AD-A080 709 WISCONSIN UNIV-MADISON DEPT OF STATISTICS F/G 12/1
AN INTRODUCTION TO APPLIED MULTIPLE TIME SERIES ANALYSIS. (U)
OCT 79 G C TIAO, G E BOX DAAG29-78-G-0166
UNCLASSIFIED ARO-15734.1-M NL

1 OF 1
AD
A080709

END
DATE
FILMED
3-80
DDC

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD A080709

DDC FILE COPY

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER (19) 15734.1-M (18) ARO	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) An Introduction to Applied Multiple Time Series Analysis,		5. TYPE OF REPORT & PERIOD COVERED (9) Technical rept.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) G. C. Tiao, G. E. P. Box		8. CONTRACT OR GRANT NUMBER(s) DAAG29-78-G-0166 (15) KB-A01-73-00-15 31
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Wisconsin, Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (12) 41
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office, P. O. Box 12211, Research Triangle Park, NC 27709		12. REPORT DATE (11) Oct [redacted] 79
		13. NUMBER OF PAGES 75
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) LEVEL		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

DDC FEB 11 1980 E

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

multiple time series
vector autoregressive moving average models
exponential smoothing
intervention analysis
principal components
canonical analysis
transfer function
econometric models

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An approach to the modelling and analysis of multiple time series is proposed. Properties of a class of vector autoregressive moving average models are discussed. Modelling procedures consisting of tentative specification, estimation and diagnostic checking are outlined and illustrated by three real examples. Various eigenvalue-eigenvector analyses are presented and some alternative approaches to multiple time series are briefly discussed.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

80 2

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

40 243

TECHNICAL REPORT NO. 582

October 1979

AN INTRODUCTION TO APPLIED MULTIPLE TIME SERIES ANALYSIS

by

G.C. Tiao and G.E.P. Box

An Introduction to Applied Multiple Time Series Analysis

by

G.C. Tiao and G.E.P. Box
Department of Statistics
University of Wisconsin

This research was supported partially by Bureau of Census, Washington, D.C. under grant no. KO-AO1-73-00-1531; Army Research Office, Durham, N.C. under grant no. DAAG29-78-G-0166; and the ALCOA Foundation.

We wish to record our acknowledgement to M.R. Grupe, G.B. Hudak, W.R. Bell, and I. Chang who have been responsible for developing the Wisconsin Multiple Time Series (WMTS-1) Program, to S.C. Hillmer for help in the estimation algorithms, and to J. Ahtola for computing assistance.

Abstract

An approach to the modelling and analysis of multiple time series is proposed. Properties of a class of vector autoregressive moving average models are discussed. Modelling procedures consisting of tentative specification, estimation and diagnostic checking are outlined and illustrated by three real examples. Various eigenvalue-eigenvector analyses are presented and some alternative approaches to multiple time series are briefly discussed.

Some key words: Multiple time series, vector autoregressive moving average models, exponential smoothing, intervention analysis, principal components, canonical analysis, transfer function, econometric models.

1. Introduction

Business, economic, engineering and environmental data are often collected in roughly equally spaced time intervals e.g. hour, week, month or quarter. In many problems, such time series data may be available on several related variables of interest. For instance, in studying demand for telephones, one may have data on monthly telephone installations, housing starts and some index of business activity. As another example, to assess the trend in air pollution, time series data on air pollutants such as ozone or carbon monoxide, on input variables such as traffic count and speed, and on meteorological variables including inversion height, temperature, wind speed, etc. are usually collected. Reasons for analyzing and modeling these series jointly are:

(i) to understand the dynamic relationships among these series. They may be contemporaneously related, one series may lead to others or there may be *feedback relationships* among some of the series. A better understanding of such relationships could, for example, in an air pollution study, lead to the design of an appropriate control strategy to improve air quality.

(ii) to improve accuracy of forecasts. When there is information on one series contained in the historical data of another, better forecasts will result when the series are modelled jointly.

(iii) to obtain potentially better results in intervention analysis, smoothing and seasonal adjustment.

Let

$$\{Z_{1t}\},\ldots,\{Z_{kt}\},\ t = 0,\pm1,\pm2,\ldots \tag{1.1}$$

be k series taken in equally spaced time intervals. Writing

$$\underset{\sim}{Z}_t = (Z_{1t},\ldots,Z_{kt})' \tag{1.2}$$

we shall refer to the k series as a k-dimensional vector or multiple time series. This report sketches an approach to the modelling and analysis of $\{\underset{\sim}{Z}_t\}$. Secton 2 presents a short review of the widely used univariate (k=1) time series and transfer function models developed in [7]. Section 3 introduces a class of vector autoregressive moving average models. Model building procedures for multiple time series are discussed in Section 4 and applied to three actual examples in Section 5. Various canonical and principal component techniques are presented in Section 6. Finally, some alternative approaches to multiple time series modelling are briefly discussed in Section 7.

2. Univariate Time Series and Transfer Function Models

2.1 Univariate stochastic difference equation models

We shall write $\underset{\sim}{Z}_t = Z_t$ for k = 1 in (1.2). An important class of models for discrete univariate series originally proposed by Yule [18] and Slutsky [16], and developed by such authors as Bartlett, Kendall, Walker, Wold and Yaglom are stochastic difference equations of the form

$$\varphi_p(B)z_t = \theta_q(B)a_t \tag{2.1}$$

where $\varphi_p(B) = 1-\varphi_1B-\ldots-\varphi_pB^p$ and $\theta_q(B) = 1-\theta_1B-\ldots-\theta_qB^q$. In (2.1) the

a_t's are independently identically and normally distributed random shocks (or white noise) with zero mean and variance σ^2, B is the back shift operator such that $BZ_t = Z_{t-1}$ and $z_t = Z_t - \eta$ is the deviation of the observation Z_t from some convenient location η. A representationally useful class of nonstationary models have d zeros of $\varphi_p(B)$ on and $p_1 = p-d$ zeros outside the unit circle and all the zeros of $\theta_q(B)$ outside the unit circle where typically p, d and q are small numbers. For such models letting $\varphi_p(B) = \phi_{p_1}(B)\alpha_d(B)$ we have

$$\phi_{p_1}(B)\alpha_d(B)z_t = \theta_q(B)a_t. \tag{2.2}$$

In the case $\alpha_d(B) = (1-B)^d$, (2.2) is known as the autoregressive integrated moving average (ARIMA) process of order (p_1,d,q). By writing $\alpha_d(B)z_t = w_t$, (2.2) is reduced to the stationary ARMA process of order (p_1,q) for w_t,

$$\phi_{p_1}(B)w_t = \theta_q(B)a_t. \tag{2.3}$$

<u>The π weights and the ψ weights</u>

Assuming that the series starts at some remote past point of time, and as shown in [6], the models (2.1) may also be written

$$z_{t+\ell} = \hat{z}_t(\ell) + e_t(\ell) \tag{2.4}$$

where

$$\hat{z}_t(\ell) = \sum_{j=1}^{\infty} \pi_j^{(\ell)} z_{t-j+1} \text{ and } e_t(\ell) = \sum_{i=0}^{\ell-1} \psi_i a_{t+\ell-i}.$$

In (2.4) $\psi_0 = 1$ and the ψ_i's may be obtained by equating coefficients in

$$\varphi_p(B)\psi(B) = \theta_q(B),\ \psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots \ . \tag{2.5}$$

Also, with $\pi_j^{(1)} = \pi_j$ the π_j's may be similarly obtained from

$$\theta_q(B)\pi(B) = \varphi_p(B),\ \pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots$$

and then

$$\pi_j^{(\ell)} = \pi_{j+1}^{(\ell-1)} + \psi_{\ell-1}\pi_j. \tag{2.6}$$

<u>Forecasting future observations</u>

Now from (2.4) the conditional distribution of $z_{t+\ell}$ at some time origin t is normal with mean $\hat{z}_t(\ell)$ and variance $\sigma^2(\ell) = \sum_{i=0}^{\ell-1}\psi_i^2\sigma^2$. We shall refer to $\hat{z}_t(\ell)$ as the <u>forecast</u> at <u>origin</u> t with <u>lead</u> ℓ and $e_t(\ell)$ the corresponding forecast error. Also, regarded as a function of ℓ, $\hat{z}_t(\ell)$ will be called the forecast function at origin t.

One rationalization of the model as defined in (2.1) is as follows. Among all linear models of the form

$$\pi(B)z_t = a_t \tag{2.7}$$

the parsimonious subset of the form of (2.2) has the sensible property

(1) that the weights $\{\pi_j^{(\ell)}\}$ converge, and

(2) that for lead time $\ell > q-p$ the forecast function $\hat{z}_t(\ell)$ comes from a very rich class of smooth functions, namely mixtures of polynomials, non-explosive exponentials and cosine functions.

<u>Some simple examples</u>

(1) If in (2.2) $p_1 = 0$, $q = 1$ and $d = 0$, we have a stationary moving average model of order 1, MA(1),

$$z_t = (1-\theta B)a_t \tag{2.8}$$

In this case, z_t is a linear combination of the current shock a_t and the previous shock a_{t-1}. Here, $\psi_1 = -\theta$ and $\psi_j = 0$ for $j > 1$. Also, $\pi_j = -\theta^j$ for $j \geq 1$.

(ii) When $p_1 = 1$, $q = 0$ and $d = 0$, we have a stationary autoregressive model of order 1, AR(1)

$$(1-\phi B)z_t = a_t \tag{2.9}$$

In this case, $\psi_j = \phi^j$ for $j \geq 1$. Also, $\pi_1 = \phi$ and $\pi_j = 0$ for $j > 1$.

(iii) A model which has been widely used to represent nonstationary series is obtained by setting $p_1 = 0$, $q = 1$, and $\alpha_d(B) = 1 - B$,

$$(1-B)z_t = (1-\theta B)a_t, \tag{2.10}$$

i.e. the first difference $w_t = (1-B)z_t$ follows an MA(1) model. Here $\psi_j = 1 - \theta$ and $\pi_j = (1-\theta)\theta^{j-1}$ for $j \geq 1$. In this case, $\hat{z}_t(\ell)$ in (2.4) becomes

$$\hat{z}_t(\ell) = (1-\theta)\sum_{j=1}^{\infty}\theta^{j-1}z_{t+1-j}, \quad (1-\theta)\sum_{j=1}^{\infty}\theta^{j-1} = 1, \tag{2.11}$$

so that forecast of all future observations is an exponentially decaying weighted average of current and past observations; this is commonly known as the method of "exponential smoothing".

<u>Multiplicative seasonal models</u>

By introducing terms of B^s in (2.2) is it possible to obtain parsimonious models for a seasonal time series of period s. A useful type of model takes the multiplicative form

$$\phi_{p_s}(B^s)\phi_{p_1}(B)\alpha_d(B)z_t = \theta_{q_s}(B^s)\theta_q(B)a_t \tag{2.12}$$

where $\phi_{p_s}(B^s)$ and $\theta_{q_s}(B^s)$ are polynomials in B^s analogous to $\phi_{p_1}(B)$ and $\theta_q(B)$ respectively. As an example, for $s = 12$, $p_s = p_1 = 0$, $q_s = q = 1$ and $\alpha_d(B) = (1-B)(1-B^{12})$ we have

$$(1-B^{12})(1-B)z_t = (1-\theta_{12}B^{12})(1-\theta B)a_t, \tag{2.13}$$

which has been widely used to represent many seasonal monthly time series.

<u>Model building procedure</u>

In [7] an iterative procedure for building models of the form of (2.2) was proposed involving

(a) <u>Specification</u> (Identification) - tentative choice of p_1, d, q by study of the sample autocorrelation function (ACF) and sample partial autocorrelation function (PACF),

(b) <u>Estimation</u> - of $\underset{\sim}{\phi} = (\phi_1,\ldots,\phi_p)'$ and $\underset{\sim}{\theta} = (\theta_1,\ldots,\theta_q)'$ by maximizing the likelihood,

(c) <u>Diagnostic checking</u> - criticism of the fitted model by study of residuals.

In some contexts a linear model of the form of (2.2) more appropriately describes the behavior of some parametric transformation $z^{(\underset{\sim}{\lambda})}$ of z, where $\underset{\sim}{\lambda}$ is a set of transformation parameters. The specific transformation may in some cases be suggested by the problem context; in others its parameter(s) may be estimated by maximum likelihood as for example is described in [4].

2.2 Transfer function models

When k series $\{z_{1t}\},\ldots,\{z_{kt}\}$ are of interest relationships sometime exist which can be represented by linear transfer function models of the form

$$z_{ht} = \sum_{i=1}^{h-1} \frac{\omega_{s_i}(B)B^{b_i}}{\delta_{r_i}(B)} z_{it} + \frac{\theta_{q_h}(B)}{\varphi_{p_h}(B)} a_{ht}, \ (h=1,2,\ldots,k) \qquad (2.14)$$

with $z_{0t} \equiv 0$, where $\omega_{s_i}(B)$, $\delta_{r_i}(B)$, $\varphi_{p_h}(B)$ and $\theta_{q_h}(B)$ are polynomials in B, the b_i's are nonnegative integers, and $\{a_{1t}\},\ldots\{a_{kt}\}$ are k independent Gaussian white noise processes with zero means and variances $\sigma_1^2,\ldots,\sigma_k^2$. A discussion of the building of elementary transfer function models was given in [7]. Also intervention models of this form with one or more of the z_h's being indicator variables have proved useful in environmental and other problems, [9].

A transfer function model may sometimes be used to relate some series z_2 of interest to a leading indicator z_1 (or a number of such indicators). This (i) can throw light on the dynamic relationship, if any, between z_1 and z_2 and (ii) depending on the ability of the leading indicator z_1 to supply additional information not already supplied by the past of z_2, can result in better forecasts of z_2.

Transfer function models of the form (2.14) assume a triangular relationship between the time series. That is to say that there is some way of arranging the series such that in addition to its own past, z_2 depends only on the present and past of z_1; z_3 on the present and past of z_2 and z_1 and so on.

However if, for example, not only does z_1 depend on the past of z_2, but z_2 depends on the past of z_1, then we must have a model which allows for this feedback.

3. Multiple Stochastic Difference Equation Models

3.1 The vector ARMA model

More general multiple time series models allow such feedback to be taken into account. These models are generalizatons of (2.1). For k series $\{\underset{\sim}{Z}_t\}$, the vector ARMA model takes the form

$$\underset{\sim}{\varphi}_p(B)\underset{\sim}{z}_t = \underset{\sim}{\theta}_q(B)\underset{\sim}{a}_t \qquad (3.1)$$

where

$$\underset{\sim}{\varphi}_p(B) = \underset{\sim}{I} - \underset{\sim}{\varphi}_1 B - \cdots - \underset{\sim}{\varphi}_p B^p, \quad \underset{\sim}{\theta}_q(B) = \underset{\sim}{I} - \underset{\sim}{\theta}_1 B - \cdots - \underset{\sim}{\theta}_q B^q$$

are matrix polynomials in B, the $\underset{\sim}{\varphi}$'s and $\underset{\sim}{\theta}$'s are k×k matrices, $\underset{\sim}{z}_t = \underset{\sim}{Z}_t - \underset{\sim}{\eta}$ is the vector of deviations from some origin $\underset{\sim}{\eta}$ which is the mean if the series is stationary, and $\{\underset{\sim}{a}_t\}$ with $\underset{\sim}{a}_t = (a_{1t},\ldots,a_{kt})'$ is a sequence of random shock vectors identically independently and normally distributed with zero mean and covariance matrix $\underset{\sim}{\Sigma}$. We shall suppose that the zeros of the determinantal polynomials $|\underset{\sim}{\varphi}_p(B)|$ and $|\underset{\sim}{\theta}_q(B)|$ are on or outside the unit circle. The series $\underset{\sim}{z}_t$ will be stationary when the zeros of $|\underset{\sim}{\varphi}_p(B)|$ are all outside the unit circle, and will be invertible when those of $|\underset{\sim}{\theta}_q(B)|$ are all outside the unit circle. Properties of the model have been discussed in [13] and [15].

The $\underset{\sim}{\pi}$ and $\underset{\sim}{\psi}$ weights

Paralleling the results in (2.4), when the series is invertible, we can write

$$\underset{\sim}{z}_{t+\ell} = \hat{\underset{\sim}{z}}_t(\ell) + \underset{\sim}{e}_t(\ell) \qquad (3.2)$$

where

$$\hat{z}_t(\ell) = \sum_{j=1}^{\infty} \pi_j^{(\ell)} z_{t+1-j}, \quad e_t(\ell) = \sum_{i=0}^{\ell-1} \psi_i a_{t+\ell-i} \, .$$

The $\pi_j^{(\ell)}$'s and ψ_j's are k×k matrices obtained from the relations

$$\varphi_p(B)\psi(B) = \theta_q(B), \quad \psi(B) = \psi_0 + \psi_1 B + \psi_2 B^2 + \ldots, \quad \psi_0 = I$$
$$\theta_q(B)\pi(B) = \varphi_p(B), \quad \pi(B) = I - \pi_1 B - \pi_2 B^2 - \ldots,$$

and

$$\pi_j^{(\ell)} = \pi_{j+1}^{(\ell-1)} + \psi_{\ell-1}\pi_j \text{ with } \pi_j^{(1)} = \pi_j .$$

<u>Forecasting future observations</u>

Similarly, in (3.2) $\hat{z}_t(\ell)$ is the conditional expectation of $z_{t+\ell}$

$$\hat{z}_t(\ell) = E(z_{t+\ell} | z_t, z_{t-1}, \ldots)$$

and (3.3)

$$e_t(\ell) = z_{t+\ell} - \hat{z}_t(\ell), \quad \ell = 1, \ldots$$

is the vector of forecast errors made at origin t. In practice, the forecast vector $\hat{z}_t(\ell)$ is obtained recursively using the formula

$$\hat{z}_t(\ell) = \varphi_1 \underset{t}{E}(z_{t+\ell-1}) + \ldots + \varphi_p \underset{t}{E}(z_{t+\ell-p}) - \theta_1 \underset{t}{E}(a_{t+\ell-1}) - \ldots - \theta_q \underset{t}{E}(a_{t+\ell-q})$$

where (3.3a)

$$\underset{t}{E}(z_{t+j}) = \begin{cases} \hat{z}_t(j) & j \geq 1 \\ z_{t+j} & \text{otherwise} \end{cases} \text{ and } \underset{t}{E}(a_{t+j}) = \begin{cases} 0 & j \geq 1 \\ a_{t+j} & \text{otherwise} \end{cases} .$$

The error vector $e_t(\ell)$ is normally distributed with zero mean and covariance matrix

$$\mathrm{Cov}(e_t(\ell)) = \sum_{i=0}^{\ell-1} \psi_i \Sigma \psi_i' \, . \quad (3.4)$$

<u>Some special cases</u>

(i) As a first example, consider the simple stationary vector MA(1) model (p=0,q=1),

$$z_t = (I - \theta B) a_t \, . \quad (3.5)$$

For k = 2, writing

$$\theta = \begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix}$$

we have that

$$\begin{aligned} z_{1t} &= a_{1t} - \theta_{11} a_{1(t-1)} - \theta_{12} a_{2(t-1)} \\ z_{2t} &= a_{2t} - \theta_{21} a_{1(t-1)} - \theta_{22} a_{2(t-1)} \end{aligned} \quad (3.6)$$

Thus, the z_{jt} depends only on the current shock a_{jt} and the elements of the shock vector one period ago, a_{t-1}. It is easy to show that individually, each series follows a univariate MA(1) model, i.e.

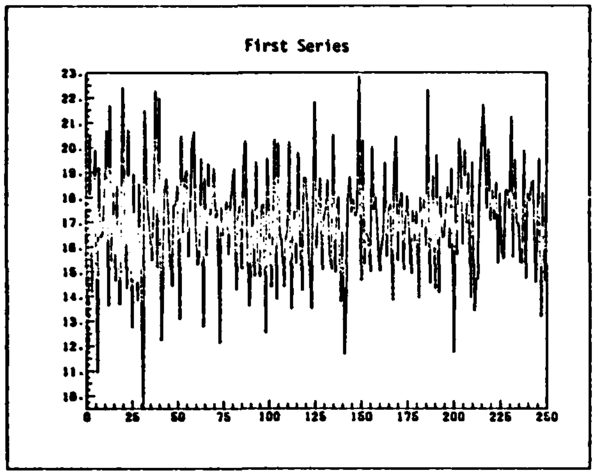
$$z_{jt} = (1 - \delta_j B) c_{jt}, \quad j = 1,2 \quad (3.7)$$

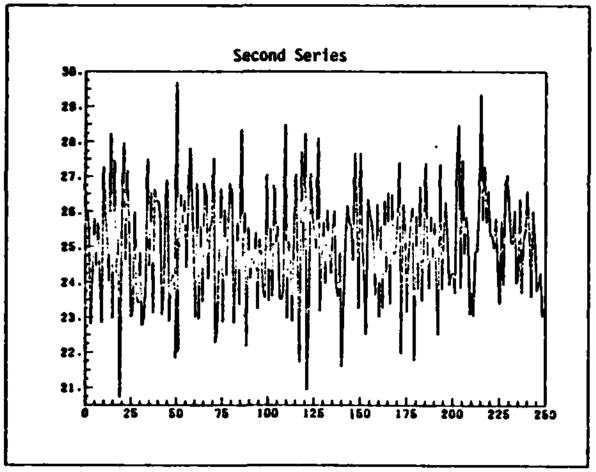
where $\{c_{jt}\}$ is some Gaussian white noise process.

Figure 3.1 shows two series with 250 observations generated from the model in (3.5) with

$$\theta = \begin{bmatrix} .2 & .3 \\ -.6 & 1.1 \end{bmatrix} \text{ and } \Sigma = \begin{bmatrix} 4 & 1 \\ 1 & 1 \end{bmatrix} . \quad (3.8)$$

Figure 3.1 Data generated from a bivariate MA(1) model in (3.5) with parameter values in (3.8).





(ii) As a second example, we consider the vector AR(1) model ($p=1, q=0$),

$$(\underset{\sim}{I}-\underset{\sim}{\varphi}B)\underset{\sim}{z}_t = \underset{\sim}{a}_t \tag{3.9}$$

Again for $k = 2$, with

$$\underset{\sim}{\varphi} = \begin{bmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{bmatrix}$$

we can write the model as

$$\begin{aligned} z_{1t} &= \varphi_{11} z_{1(t-1)} + \varphi_{12} z_{2(t-1)} + a_{1t} \\ z_{2t} &= \varphi_{21} z_{1(t-1)} + \varphi_{22} z_{2(t-1)} + a_{2t} \end{aligned} \tag{3.10}$$

If we regard z_{1t} and z_{2t} as dependent variables and $z_{1(t-1)}$ and $z_{2(t-1)}$ as input or independent variables, then (3.10) is in the form of a bivariate linear model, a point which will be discussed later in further detail. Now multiplying both sides of (3.9) by the adjoint of $(\underset{\sim}{I}-\underset{\sim}{\varphi}B)$, we obtain, for $k = 2$,

$$[(1-\varphi_{11}B)(1-\varphi_{22}B)-\varphi_{12}\varphi_{21}B^2]\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} 1-\varphi_{22}B & -\varphi_{12}B \\ -\varphi_{21}B & 1-\varphi_{11}B \end{bmatrix}\begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} \tag{3.11}$$

Thus, each series individually follows an ARMA (2,1) model. Note, however, that this is the maximum order for each individual series and that the autoregressive parts need not be identical as would appear from (3.11). For example, suppose $\varphi_{12} = \varphi_{21} = 0$, then each would follow an AR(1) model.

Figure 3.2 shows two series with 150 observations generated from (3.9) with

$$\underset{\sim}{\varphi} = \begin{bmatrix} .2 & .3 \\ -.6 & 1.1 \end{bmatrix} \text{ and } \underset{\sim}{\Sigma} = \begin{bmatrix} 4 & 1 \\ 1 & 1 \end{bmatrix} \tag{3.12}$$

(iii) As a third example, consider the model

$$(1-B)\underset{\sim}{z}_t = (\underset{\sim}{I}-\underset{\sim}{\theta}B)\underset{\sim}{a}_t \tag{3.13}$$

i.e. after differencing each series we obtain a vector MA(1) model. In this case $\underset{\sim}{\psi}_j = \underset{\sim}{I}-\underset{\sim}{\theta}$ and, when the series is invertible, $\underset{\sim}{\pi}_j = (\underset{\sim}{I}-\underset{\sim}{\theta})\underset{\sim}{\theta}^{j-1}$, $j \geq 1$. Thus, the forecast $\hat{\underset{\sim}{z}}_t(\ell)$ takes the same form as in (2.11)

$$\hat{\underset{\sim}{z}}_t(\ell) = (\underset{\sim}{I}-\underset{\sim}{\theta})\sum_{j=1}^{\infty}\underset{\sim}{\theta}^{j-1}\underset{\sim}{z}_{t+1-j} \tag{3.14}$$

except that the weights are now matrices. Since

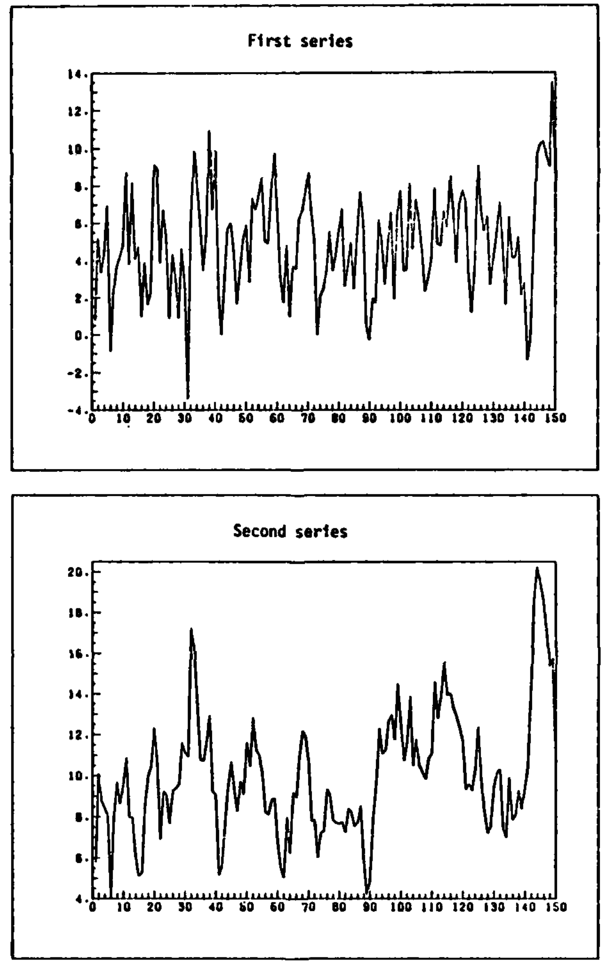
$$(\underset{\sim}{I}-\underset{\sim}{\theta})\sum_{j=1}^{\infty}\underset{\sim}{\theta}^{j-1} = \sum_{j=1}^{\infty}\underset{\sim}{\pi}_j = \underset{\sim}{I}$$

it follows that each element of $\hat{\underset{\sim}{z}}_t(\ell)$ is a weighted linear combination of the elements of $\underset{\sim}{z}_t, \underset{\sim}{z}_{t-1}, \ldots$ with weights summing to one. Figure 3.3 shows the weight functions for k = 2

$$\underset{\sim}{\pi}_j = \begin{bmatrix} \pi_{j11} & \pi_{j12} \\ \pi_{j21} & \pi_{j22} \end{bmatrix} \text{ with } \underset{\sim}{\theta} = \begin{bmatrix} .2 & .3 \\ -.6 & 1.1 \end{bmatrix}. \tag{3.15}$$

It can in fact be shown that the elements of $\underset{\sim}{\pi}_j$ when regarded as functions of j are mixtures of exponentials and cosine functions. Finally it is readily shown that for the model in (3.13), each series z_{jt} individually follows a univariate exponential smoothing model of the form (2.10).

Figure 3.2 Data generated from a bivariate AR(1) model in (3.9) with parameter values in (3.12).



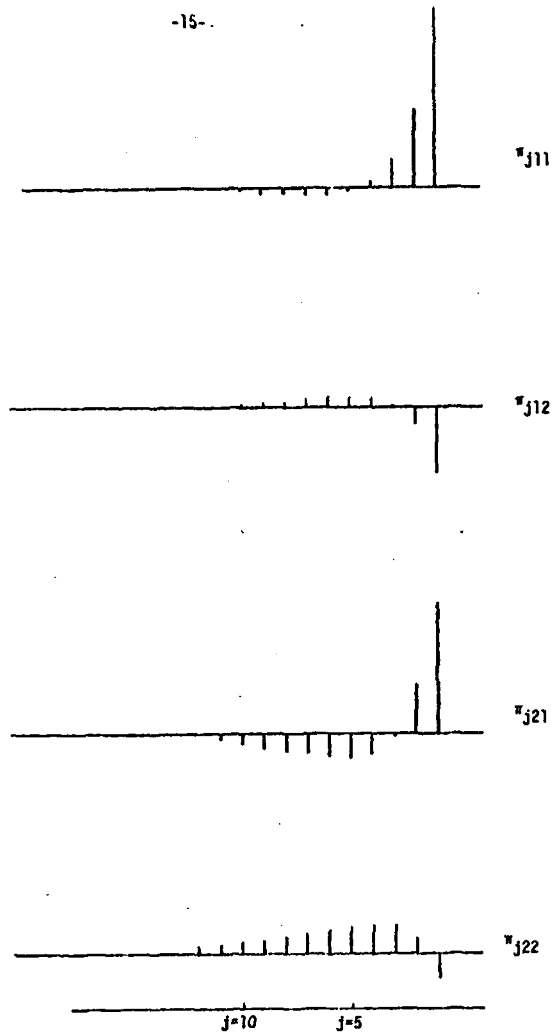


Figure 3.3 Elements of $\underset{\sim}{\pi}_j$ for the Bivariate Exponential Smoothing Model (3.13) with parameter values in (3.15).

3.2 Vector ARMA model and transfer function model

For the vector model in (3.1), in general, all elements of $\underset{\sim}{z}_t$ are related to all elements of $\underset{\sim}{z}_{t-j}$ $(j=1,2,\ldots)$ and there can be feedback relationships between all the series. However, if the $\underset{\sim}{z}_t$'s can be arranged so that the coefficient matrices $\underset{\sim}{\varphi}$'s and $\underset{\sim}{\theta}$'s are all lower triangular then so will be all the $\underset{\sim}{\pi}_j$ matrices and (3.1) can be written as a transfer function model of the form (2.14).

To illustrate, for the MA(1) model in (3.5) with $k = 2$, suppose $\underset{\sim}{\theta}$ is lower triangular, so that

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} (1-\theta_{11}B) & \\ -\theta_{21}B & (1-\theta_{22}B) \end{bmatrix} \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} \tag{3.16}$$

Writing $a_{2t} = \beta a_{1t} + \epsilon_t$ where ϵ_t and a_{1t} are independent, we can express (3.16) in the alternate form

$$\begin{aligned} z_{1t} &= (1-\theta_{11}B)a_{1t} \\ z_{2t} &= \frac{\omega_0-\omega_1 B}{1-\theta_{11}B} z_{1t} + (1-\theta_{22}B)\epsilon_t \end{aligned} \tag{3.16a}$$

where $\omega_0 = \beta$ and $\omega_1 = \beta\theta_{22} + \theta_{21}$, which is a special case of the transfer function model (2.14).

The transfer function form (2.14) can be generalized in a multiple time series setting by allowing the $\underset{\sim}{\varphi}$'s and $\underset{\sim}{\theta}$'s to be all lower block triangular. As an illustration, suppose

$$\begin{bmatrix} \underset{\sim}{z}_{1t} \\ \underset{\sim}{z}_{2t} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{I}-\underset{\sim}{\theta}_{11}B & \\ -\underset{\sim}{\theta}_{21}B & \underset{\sim}{I}-\underset{\sim}{\theta}_{22}B \end{bmatrix} \begin{bmatrix} \underset{\sim}{a}_{1t} \\ \underset{\sim}{a}_{2t} \end{bmatrix} \tag{3.17}$$

where z_{1t} and a_{1t} are $k_1 \times k_1$ vectors, z_{2t} and a_{2t} are $k_2 \times k_2$ vectors and θ_{11}, θ_{21} and θ_{22} are matrices of appropriate dimensions. Thus, the input series $\{z_{1t}\}$, as well as the output series $\{z_{2t}\}$, themselves are allowed to have feedback relationships.

3.3 Relationship to structural equations in econometrics

In econometric literature, one often encounters linear structural equations of the form

$$A(B)Y_t + H(B)X_t = U_t \tag{3.18}$$

where Y_t is the $k_1 \times 1$ endogenous vector, X_t the $k_2 \times 1$ exogenous vector, U_t the $k_1 \times 1$ error vector, $A(B) = A_0 + A_1 B + \ldots + A_r B^r$ and $H(B) = H_0 + H_1 B + \ldots + H_s B^s$ are, respectively, $k_1 \times k_1$ and $k_1 \times k_2$ matrix polynomials, and U_t and X_t are independent. This can be written as

$$Y_t - A_1^* Y_{t-1} - \cdots - A_r^* Y_{t-r} + H_0^* X_t + \cdots + H_s^* X_{t-s} = \varepsilon_t \tag{3.19}$$

where $A_j^* = A_0^{-1} A_j$, $H_\ell^* = A_0^{-1} H_\ell$ and $\varepsilon_t = A_0^{-1} U_t$. If, in addition, we suppose that X_t follows an ARMA model, say

$$X_t - G_1 X_{t-1} - \cdots - G_{p*} X_{t-p*} = b_t - M_1 b_{t-1} - \cdots - M_{q*} b_{t-q*} \tag{3.20}$$

where the G's and the M's are $k_2 \times k_2$ matrices, and $\{b_t\}$ a sequence of random shock vectors, then writing $z_t' = [X_t', Y_t']$, it is clear that the model for z_t will be of the ARMA type with lower block triangular φ's and θ's and, therefore, be in the transfer function form. We may thus regard (3.1) as a "reduced form" of (3.18) and (3.20). We shall discuss this point further in Section 7.

3.4 Cross covariance and correlation matrices

When the multiple time series $\{Z_t\}$ is stationary, with mean vector η, then the lag ℓ cross-covariance matrix is defined as the expectation

$$E(z_{t-\ell} z_t') = \Gamma(\ell) = \{\gamma_{ij}(\ell)\}, \quad \begin{matrix} \ell = 0, \pm 1, \pm 2, \ldots \\ i, j = 1, \ldots, k \end{matrix} \tag{3.21}$$

and the corresponding cross correlation matrix is

$$\rho(\ell) = \{\rho_{ij}(\ell)\} \tag{3.22}$$

where

$$\rho_{ij}(\ell) = \gamma_{ij}(\ell)/[\gamma_{ii}(0)\gamma_{jj}(0)]^{1/2} .$$

Note that $\Gamma(-\ell) = \Gamma'(\ell)$ and $\rho(-\ell) = \rho'(\ell)$.

For the vector ARMA model in (3.1), assuming stationarity we can write $z_t = \varphi_p^{-1}(B)\theta_q(B)a_t = \psi(B)a_t$ so that

$$E(z_t a_{t-j}') = \begin{cases} \Sigma & j = 0 \\ \psi_j \Sigma & j \geq 1 \\ 0 & j < 0 \end{cases} . \tag{3.23}$$

Writing

$$z_{t-\ell}(z_t' - z_{t-1}'\varphi_1' - \cdots - z_{t-p}'\varphi_p') = z_{t-\ell}(a_t' - a_{t-1}'\theta_1' - \cdots - a_{t-q}'\theta_q')$$

and taking expectation on both sides of the equation, we obtain

$$\Gamma(\ell) = \begin{cases} \sum_{j=\ell-r}^{\ell-1} \Gamma(j)\varphi_{\ell-j}' - \sum_{j=0}^{r-\ell} \psi_j \Sigma \theta_{j+\ell}', & \ell = 0, \ldots, r \\ \sum_{j=1}^{r} \Gamma(\ell-j)\varphi_j', & \ell > r \end{cases} \tag{3.24}$$

where $\underset{\sim}{\theta}_0 = -\underset{\sim}{I}$, $r = \max(p,q)$ and it is understood that (i) if $p < q$, $\underset{\sim}{\varphi}_{p+1} = \cdots = \underset{\sim}{\varphi}_r = 0$, and (ii) if $q < p$, $\underset{\sim}{\theta}_{q+1} = \cdots = \underset{\sim}{\theta}_r = \underset{\sim}{0}$.

In particular when $p = 0$, i.e. we have a vector MA(q) model, then

$$\underset{\sim}{\Gamma}(\ell) = \begin{cases} \sum_{j=0}^{q-\ell} \underset{\sim}{\theta}_j \underset{\sim}{\Sigma} \underset{\sim}{\theta}'_{j+\ell}, & \ell = 0,\ldots,q \\ \underset{\sim}{0} & \ell > q \end{cases} \quad . \tag{3.25}$$

Thus, all auto and cross correlations will be zero when $\ell > q$.

<u>Two examples</u>

For the MA(1) model (3.5) with parameter values given in (3.8), we find

$$\underset{\sim}{\Gamma}(0) = \begin{bmatrix} 4.37 & .89 \\ .89 & 2.33 \end{bmatrix}, \ \underset{\sim}{\Gamma}(1) = -\begin{bmatrix} 1.1 & -1.3 \\ .5 & .5 \end{bmatrix}$$

so that

$$\underset{\sim}{\rho}(1) = \begin{bmatrix} -.25 & .41 \\ -.16 & -.21 \end{bmatrix} \tag{3.26}$$

and $\underset{\sim}{\rho}(\ell) = \underset{\sim}{0}$ for $\ell > 1$.

For the AR(1) model in (3.9), we have from (3.24)

$$\begin{cases} \underset{\sim}{\Gamma}(0) = \underset{\sim}{\varphi}\underset{\sim}{\Gamma}(0)\underset{\sim}{\varphi}' + \underset{\sim}{\Sigma} \\ \underset{\sim}{\Gamma}(\ell) = \underset{\sim}{\Gamma}(\ell-1)\underset{\sim}{\varphi}', \ \ell > 1 \end{cases} \quad . \tag{3.27}$$

Note that for given $\underset{\sim}{\varphi}$ and $\underset{\sim}{\Sigma}$, the first equation is linear in the elements of $\underset{\sim}{\Gamma}(0)$. For k = 2 and the parameter values in (3.12), Figures 3.3(a)-(c) show the auto and cross correlations $\rho_{11}(\ell)$, $\rho_{22}(\ell)$ and $\rho_{12}(\ell)$, respectively. Unlike the moving average case, the correlations decay gradually to zero as $|\ell|$ increases.

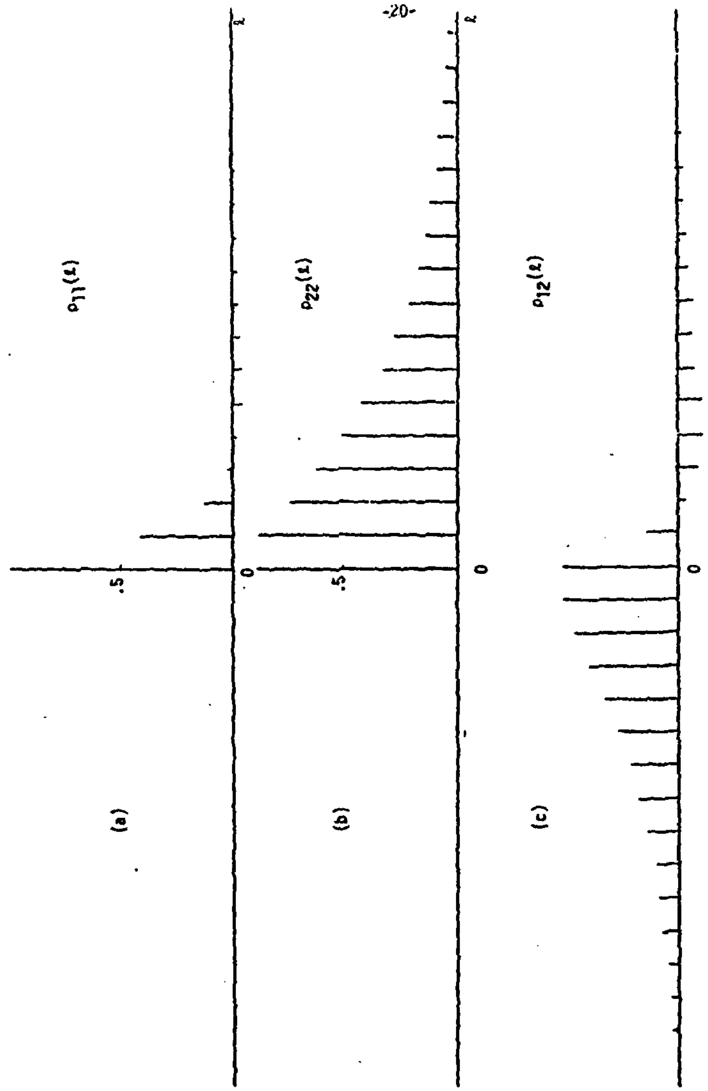


Figure 3.3 Auto and Cross Correlations of a Bivariate AR(1) Model with Parameter Values in (3.12)

3.5 Partial correlation matrices

Analogous to the partial autocorrelation function in the univariate case, see e.g. [7], we may define a generalized partial cross correlation matrix function $\underset{\sim}{P}(\ell)$ in the following manner. If the series $\{\underset{\sim}{z}_t\}$ follows a stationary AR(ℓ) model, we require that

$$\underset{\sim}{P}(\ell) = \underset{\sim}{\Phi}_\ell, \ \ell = 1,2,\ldots \tag{3.28}$$

From (3.24), and setting $\underset{\sim}{\theta}_j = \underset{\sim}{0}$ for $j \geq 1$, we define $\underset{\sim}{P}(\ell)$ in terms of the cross covariance matrices $\underset{\sim}{\Gamma}(\ell)$'s as

$$\underset{\sim}{P}'(\ell) = \begin{cases} \underset{\sim}{\Gamma}^{-1}(0)\underset{\sim}{\Gamma}(1), & \ell = 1 \\ (\underset{\sim}{\Gamma}(0)-\underset{\sim}{b}'_\ell \underset{\sim}{A}_\ell^{-1}\underset{\sim}{b}_\ell)^{-1}[\underset{\sim}{\Gamma}(\ell)-\underset{\sim}{b}'_\ell \underset{\sim}{A}_\ell^{-1}\underset{\sim}{c}_\ell], & \ell > 1 \end{cases} \tag{3.29}$$

where

$$\underset{\sim}{A}_\ell = \begin{bmatrix} \underset{\sim}{\Gamma}(0) & \cdots & \underset{\sim}{\Gamma}'(\ell-2) \\ \vdots & \ddots & \vdots \\ \underset{\sim}{\Gamma}(\ell-2) & \cdots & \underset{\sim}{\Gamma}(0) \end{bmatrix} \quad \underset{\sim}{b}_\ell = \begin{bmatrix} \underset{\sim}{\Gamma}'(\ell-1) \\ \vdots \\ \underset{\sim}{\Gamma}'(1) \end{bmatrix} \quad \underset{\sim}{c}_\ell = \begin{bmatrix} \underset{\sim}{\Gamma}(1) \\ \vdots \\ \underset{\sim}{\Gamma}(\ell-1) \end{bmatrix}.$$

It follows from this definition that if $\underset{\sim}{z}_t$ follows an AR(p) model, then

$$\underset{\sim}{P}(\ell) = \underset{\sim}{0} \text{ for } \ell > p. \tag{3.30}$$

4. Model Building Strategy for Multiple Time Series

The class of vector autoregressive moving average models (3.1) is very extensive. Given data in the form of a time series of length n, we attempt to find a model which, while containing as few parameters as possible, represents adequately the dependencies in the data at hand. Extending the model building strategy developed in [7], a preliminary version of a computer package for the analysis of multiple time series has been completed. The package consists of three main programs: (i) Preliminary Analysis, (ii) Stepwise Autoregression and (iii) Estimation and Forecasting.

The process of model selection is iterative in nature and consists of three main stages: (i) tentative specification (identification), (ii) estimation, and (iii) diagnostic checking. We now give a comprehensive description of the principal tools used in our modelling approach.

4.1 Tentative specification

The aim here is to employ statistics which (a) can be readily calculated from the data and (b) allow the experimenter to select a subclass of models for further consideration. Specific methods employed are:

(i) sample cross correlation matrices $\hat{\underset{\sim}{\rho}}(\ell) = \{\hat{\rho}_{ij}(\ell)\}$ of the original data where

$$\{\hat{\rho}_{ij}(\ell)\} = \sum(Z_{it}-\bar{Z}_i)(Z_{j(t+\ell)}-\bar{Z}_j)/(\sum(Z_{it}-\bar{Z}_i)^2\sum(Z_{jt}-\bar{Z}_j)^2)^{1/2} \tag{4.1}$$

where $\bar{Z}_i$ is the sample mean of the i^{th} component series of $\underset{\sim}{Z}_t$, or the matrices of appropriately differenced data.

(ii) estimates of the generalized partial cross correlation matrices obtained by fitting successive autoregressive models of increasing order.

(iii) sample cross correlation matrices of the residual series from each fitted autoregressive model.

Sample cross correlations

The sample cross correlations $\hat{\rho}_{ij}(\ell)$'s are estimates of the theoretical values $\rho_{ij}(\ell)$'s. They are particularly useful in spotting low order vector moving average models since from (3.25) $\rho_{ij}(\ell) = 0$ for $\ell > q$.

For the data shown in Figure 3.1 generated from the bivariate MA(1) model in (3.8), Figures 4.1(a)-(c) show, respectively, the sample autocorrelations $\hat{\rho}_{11}(\ell)$ of z_{1t}, the sample autocorrelations $\hat{\rho}_{22}(\ell)$ of z_{2t}, and the sample cross correlations $\hat{\rho}_{12}(\ell)$ between these two series. As expected, large values (in magnitude) occur at $|\ell| = 1$, and one would be led to tentatively specify the model as a MA(1).

While graphs of this kind have proved useful in the analysis of one or two series, it will become increasingly cumbersome as the number of series is increased. For example, one would need to simultaneously inspect 10 graphs when k = 4 and 15 when k = 5.

For vector series, it seems more useful to list the sample cross correlation matrices $\hat{\underset{\sim}{\rho}}(\ell)$ themselves in sequence as a function of the lag ℓ. The absence of large (in magnitude) elements after a certain lag would indicate directly the order of the moving average model. For the data in Figure 3.1, the $\hat{\underset{\sim}{\rho}}(\ell)$ matrices are shown in Table 4.1(a) for $\ell = 1,\ldots,12$.

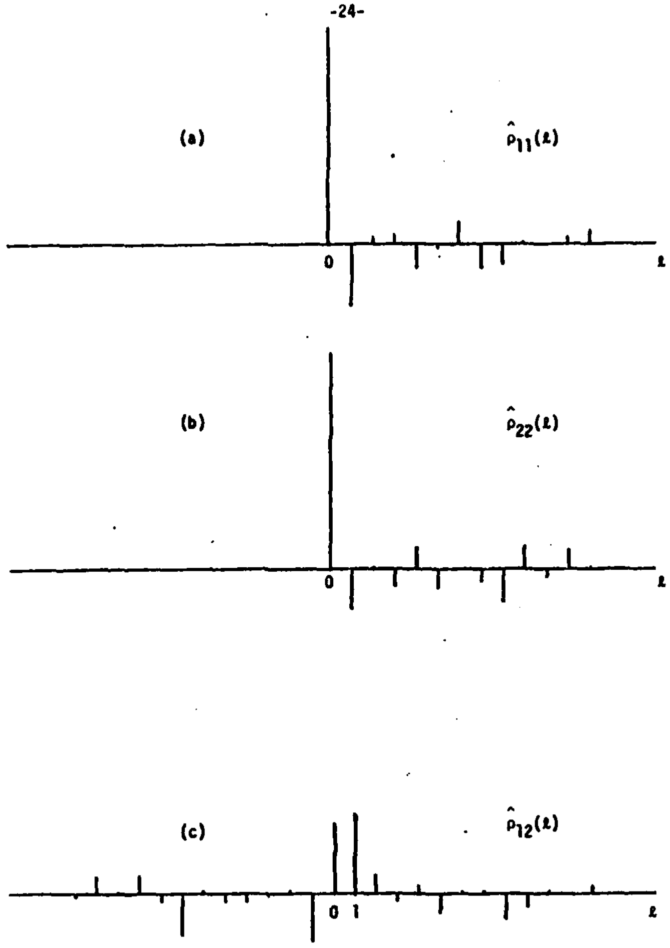


Figure 4.1 Sample Auto and Cross Correlations for the Data in Figure 3.1.

To summarize the structure of the cross correlation matrices, one may use the following device which is motivated from the consideration that if the series were white noise, then for large n, the $\hat{\rho}_{ij}(\ell)$'s would be normally distributed with mean 0 and variance n^{-1}. Thus, instead of the numerical values, a plus sign is used to indicate a value greater than $2n^{-1/2}$, a minus sign a value less than $-2n^{-1/2}$ and a dot to indicate a value in between $-2n^{-1/2}$ and $2n^{1/2}$. This simple device greatly facilitates the user's comprehension of the mass of information contained in the matrices, as is illustrated in Table 4.1(b). Another useful summary is to list the symbols for each element in the matrix over all lags as shown in Table 4.1(c).

For the series shown in Figure 3.2 generated from an AR(1) model, the sample cross correlation matrices in terms of the plus, minus and dot symbols are given in Table 4.2. As expected, significant values occur over many lag values, indicating that an autoregressive model might be appropriate.

In summary, the pattern of symbols in the sequence of cross correlation matrices makes it very easy to choose between an autoregressive or a moving average model, and for the latter to tentatively select the appropriate order.

<u>Sample generalized partial correlations and related summary statistics</u>

Sample estimates of the generalized partial cross correlation matrices $\underset{\sim}{P}(\ell)$ in (3.29) are useful in identifying the order of an autoregressive model. In our computer package, the estimates $\hat{\underset{\sim}{P}}(\ell)$ are

<u>Table 4.1</u>

(a) Sample cross correlation matrices $\hat{\underset{\sim}{\rho}}(\ell)$ for the data in Figure 3.1

Lag 1-6

$$\begin{bmatrix} -.28 & .37 \\ -.21 & -.19 \end{bmatrix} \begin{bmatrix} .03 & .08 \\ .02 & .01 \end{bmatrix} \begin{bmatrix} .04 & -.03 \\ -.01 & -.08 \end{bmatrix} \begin{bmatrix} -.11 & .04 \\ -.03 & .09 \end{bmatrix} \begin{bmatrix} -.02 & -.09 \\ -.02 & -.08 \end{bmatrix} \begin{bmatrix} .10 & .01 \\ .01 & -.00 \end{bmatrix}$$

Lag 7-12

$$\begin{bmatrix} -.11 & .01 \\ -.17 & -.06 \end{bmatrix} \begin{bmatrix} -.09 & -.12 \\ -.03 & -.16 \end{bmatrix} \begin{bmatrix} .01 & -.06 \\ .08 & .10 \end{bmatrix} \begin{bmatrix} -.00 & .02 \\ .01 & -.04 \end{bmatrix} \begin{bmatrix} .03 & .00 \\ .08 & .08 \end{bmatrix} \begin{bmatrix} .06 & .04 \\ -.01 & .01 \end{bmatrix}$$

(b) $\hat{\underset{\sim}{\rho}}(\ell)$ in term of plus, minus and dot symbols

Lag 1-6

$$\begin{bmatrix} - & + \\ - & - \end{bmatrix} \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}$$

Lag 7-12

$$\begin{bmatrix} \cdot & \cdot \\ - & \cdot \end{bmatrix} \begin{bmatrix} \cdot & \cdot \\ \cdot & - \end{bmatrix} \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}$$

(c) Pattern of correlations for each element in the matrix over all lags

	z_1	z_2
z_1	-...........	+...........
z_2	-.....-.....	-......-....

Table 4.2

Sample cross correlation matrices $\hat{\underline{\rho}}(\ell)$ for the data in Figure 3.1 in terms of plus, minus and dot symbols

Lag 1-6

$$\begin{bmatrix} + & \cdot \\ + & + \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ + & + \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ + & + \end{bmatrix}\begin{bmatrix} \cdot & - \\ + & + \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ + & + \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ + & \cdot \end{bmatrix}$$

Lag 7-12

$$\begin{bmatrix} - & - \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} - & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}$$

obtained by fitting autoregressive models of $\ell = 1,2,\ldots$ successively. Specifically, a vector AR(p) model can be written

$$\underline{z}'_t = \underline{z}'_{t-1}\underline{\Phi}'_1 + \ldots + \underline{z}'_{t-p}\underline{\Phi}'_p + \underline{a}'_t \quad . \tag{4.2}$$

Thus, if we have n observations, then ignoring the end effect, we can express the above in the form of a multivariate linear model

$$\underline{Y} = \underline{X}_1\underline{\Phi}'_1 + \ldots + \underline{X}_p\underline{\Phi}'_p + \underline{\varepsilon} \tag{4.3}$$

where

$$\underline{Y} = \begin{bmatrix} \underline{z}'_{p+1} \\ \vdots \\ \underline{z}'_n \end{bmatrix}, \ \underline{X}_1 = \begin{bmatrix} \underline{z}'_p \\ \vdots \\ \underline{z}'_{n-1} \end{bmatrix}, \ldots, \underline{X}_p = \begin{bmatrix} \underline{z}'_1 \\ \vdots \\ \underline{z}'_{n-p} \end{bmatrix} \text{and } \underline{\varepsilon} = \begin{bmatrix} \underline{a}'_{p+1} \\ \vdots \\ \underline{a}'_n \end{bmatrix}$$

and the least squared estimates $\hat{\underline{\Phi}}'_1,\ldots,\hat{\underline{\Phi}}'_p$ can be readily determined.

Let p be the largest lag of the sample partial cross correlation matrices desired, then an estimate of $\underline{P}(\ell)$ is

$$\hat{\underline{P}}(\ell) = \hat{\underline{\Phi}}_\ell \tag{4.4}$$

when an AR(ℓ) model is fitted, $\ell = 1,2,\ldots,p$.

It is well known, see e.g. [1], that for a stationary AR(ℓ) model, asymptotically the estimates $\hat{\underline{\Phi}}'_1,\ldots,\hat{\underline{\Phi}}'$ have the same distributional properties as those in the traditional multivariate linear model in which the $\underline{X}$'s are regarded as predetermined. Using standard normal linear model theory, we can compute the estimated standard errors of elements of $\hat{\underline{P}}(\ell)$, and divide the estimates by their corresponding estimated

standard errors to obtain standardized coefficients. The pattern of the sample partial cross correlations can also be summarized by assigning a plus sign when a standardized coefficient is greater than 2, a minus sign when it is less than -2 and a dot for values between -2 and 2.

To help determine tentatively the order of an autoregressive model, we may also make reference to the likelihood ratio statistic corresponding to testing the null hypothesis $\underset{\sim}{\varphi}_\ell = \underset{\sim}{0}$ against the alternative $\underset{\sim}{\varphi}_\ell \neq \underset{\sim}{0}$ when an AR(ℓ) model is fitted. Let

$$\underset{\sim}{S}(\ell) = (\underset{\sim}{Y}-\underset{\sim}{X}_1\hat{\underset{\sim}{\varphi}}_1'-\ldots-\underset{\sim}{X}_\ell\hat{\underset{\sim}{\varphi}}_\ell')'(\underset{\sim}{Y}-\underset{\sim}{X}_1\hat{\underset{\sim}{\varphi}}_1'-\ldots-\underset{\sim}{X}_\ell\hat{\underset{\sim}{\varphi}}_\ell') \tag{4.5}$$

be the matrix of residual sum of squares and cross products after fitting an AR(ℓ). Then with $\underset{\sim}{S}(0) = \underset{\sim}{Y}'\underset{\sim}{Y}$, for $\ell = 1,\ldots,p$ the likelihood ratio statistic is the ratio of the determinants

$$|\underset{\sim}{S}(\ell)|/|\underset{\sim}{S}(\ell-1)| \quad . \tag{4.6}$$

Using Bartlett's approximation [2], the statistic

$$M(\ell) = -(N-\tfrac{1}{2}-\ell\cdot k)\log_e|S(\ell)|/|S(\ell-1)| \tag{4.7}$$

is, on the null hypothesis, asymptotically distributed as χ^2 with k^2 degrees of freedom where $N = n-p-1$ is the effective number of observations, assuming that a constant term is included in the model.

In addition to the $M(\ell)$ statistic (4.7), the diagonal elements of the residual covariance matrices $\underset{\sim}{\Sigma}$ corresponding to the successive AR models may also be of interest since they show how the fit is improved as the order is increased.

For the generated series in Figure 3.2, the matrices of standardized coefficients and corresponding summary symbols, the $M(\ell)$ statistics in (4.7) and the diagonal elements of the residual covariance matrices are shown in Table 4.3 for $\ell = 1,\ldots,5$. They indicate that an AR(1) or at most an AR(2) would be adequate for the data.

The pattern of the generalized partial cross correlation matrices and related statistics are given in Table 4.4 for the series shown in Figure 3.1. Here, if we were to use an autoregressive model to represent the data, we would perhaps need one with order as high as 7. This is not surprising since the data were generated from an MA(1) model with parameter values given in (3.8). Writing $\underset{\sim}{z}_t = \underset{\sim}{\pi}_1\underset{\sim}{z}_{t-1} + \underset{\sim}{\pi}_2\underset{\sim}{z}_{t-2} + \ldots + \underset{\sim}{a}_t$, we find

$$\underset{\sim}{\pi}_1 = \begin{bmatrix} -.2 & -.3 \\ .6 & -1.1 \end{bmatrix}, \underset{\sim}{\pi}_2 = \begin{bmatrix} .14 & -.39 \\ .78 & -1.03 \end{bmatrix}, \ldots, \underset{\sim}{\pi}_6 = \begin{bmatrix} .23 & -.25 \\ .49 & -.51 \end{bmatrix}$$
$$|\underset{\sim}{\pi}_1| = .4 \quad |\underset{\sim}{\pi}_2| = .16\ , \ \ldots\ , \ |\underset{\sim}{\pi}_6| = .0041 \tag{4.8}$$

Although the determinants $|\underset{\sim}{\pi}_j|$ decrease rapidly towards zero as j increases, the elements of $\underset{\sim}{\pi}_j$ converge to zero very slowly so that many autoregressive terms would be needed to provide an adequate approximation. On the other hand, recall that the sample cross correlation pattern in Table 4.1 indicated directly that a MA(1) model would be appropriate.

In summary, the pattern of the generalized partial cross correlation matrices, the related $M(\ell)$ statistic, and the diagonal elements of the residual covariance matrix would help distinguish between moving average or autoregressive models and, for the latter, tentatively select the appropriate order.

Table 4.3

Standardized Sample Partial Cross Correlations and Related Statistics for the data in Figure 3.2

Lag ℓ	Standardized coefficients		summary symbols		$M(\ell)$ (χ_4^2)	Diagonal elements of $\tilde{\Sigma}$
1	1.70	5.79	.	+	356.96	5.30
	-16.28	32.90	-	+		1.08
2	- 1.68	1.98	.	.	7.04	5.16
	- 1.64	2.39	.	+		1.03
3	1.20	- .54	.	.	2.63	5.07
	.30	.10	.	.		1.03
4	.90	-1.25	.	.	4.38	5.01
	- .85	.76	.	.		1.02
5	.51	.10	.	.	2.42	4.95
	1.11	- .56	.	.		1.01

Table 4.4

Pattern of Sample Partial Cross Correlations and Related Statistics for Data in Figure 3.1

Lag	Pattern of $\tilde{\Phi}(\ell)$	$M(\ell)$ (χ_4^2)	$\tilde{\Sigma}$
1	- -	123.2	4.78
	+ -		1.88
2	. .	75.9	4.75
	+ -		1.43
3	+ .	35.2	4.63
	+ -		1.23
4	. .	27.5	4.63
	+ .		1.08
5	. .	16.6	4.61
	+ .		1.04
6	. .	13.5	4.53
	+ .		.98
7	. -	16.5	4.38
	+ .		.94
8	. .	8.1	4.31
	. -		.91

<u>Sample residual cross correlation matrices after AR fit</u>

After each AR(ℓ) fit, $\ell = 1,\ldots,p$, cross correlation matrices of the residuals $\hat{\underset{\sim}{a}}_t$'s, may be readily obtained. Table 4.5 shows pattern of the correlations after fitting AR(1) and AR(2) to the data in Figure 3.2, where a plus sign is used to indicate values greater than $2n^{-1/2}$, a minus sign for values less than $-2n^{-1/2}$ and a dot for in between values. Again, they clearly indicate that there is no need to go beyond an AR(2) model.

For mixed vector autoregressive moving average models in general, both the population cross correlation matrices $\underset{\sim}{\rho}(\ell)$ and the generalized partial cross correlation matrices $\underset{\sim}{P}(\ell)$ will decay only gradually toward $\underset{\sim}{0}$, making order determination using estimates of these quantities difficult in practice. In some situations, patterns in residual cross correlations after AR fit may help spot the orders.

Consider the case of a stationary ARMA(1,1) model

$$(\underset{\sim}{I}-\underset{\sim}{\varphi}B)\underset{\sim}{z}_t = (\underset{\sim}{I}-\underset{\sim}{\theta}B)\underset{\sim}{a}_t \ . \qquad (4.9)$$

If an AR(1) model is fitted to $\{\underset{\sim}{z}_t\}$, then the estimate $\hat{\underset{\sim}{\varphi}}$ would be biased. In fact, asymptotically $\hat{\underset{\sim}{\varphi}}$ converges in probability to

$$\hat{\underset{\sim}{\varphi}} \to \underset{\sim}{\varphi}_0 = \underset{\sim}{\Gamma}'(1)\underset{\sim}{\Gamma}(0)^{-1} \ . \qquad (4.10)$$

Thus, approximately, the residuals $\hat{\underset{\sim}{a}}_t = \underset{\sim}{z}_t - \underset{\sim}{\varphi}_0\underset{\sim}{z}_{t-1}$ would follow the model

$$\hat{\underset{\sim}{a}}_t = (\underset{\sim}{I}-\underset{\sim}{\varphi}_0B)(\underset{\sim}{I}-\underset{\sim}{\varphi}B)^{-1}(\underset{\sim}{I}-\underset{\sim}{\theta}B)\underset{\sim}{a}_t \ . \qquad (4.11)$$

<u>Table 4.5</u>

Pattern of Residual Cross Correlations for the Data in Figure 3.2

AR(1) Lag 1-6

$$\begin{bmatrix} \cdot & \cdot \\ \cdot & - \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} + & \cdot \\ \cdot & \cdot \end{bmatrix}$$

Lag 7-12

$$\begin{bmatrix} - & \cdot \\ - & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}$$

AR(2) Lag 1-6

$$\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} + & \cdot \\ \cdot & \cdot \end{bmatrix}$$

Lag 7-12

$$\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}$$

For $k = 1$, $\{\hat{a}_t\}$ follows an ARMA(1,2) model so that the autocorrelations of $\hat{a}_t$ are

$$\rho_{\hat{a}}(j) = \varphi\rho_{\hat{a}}(j-1),\ j > 2 \tag{4.12}$$

and $\rho_{\hat{a}}(1)$ and $\rho_{\hat{a}}(2)$ are functions of φ and θ. Table 4.6 gives values of $\rho_{\hat{a}}(1)$ and $\rho_{\hat{a}}(2)$ for various combinations of values of φ and θ. For each combination, the first value is $\rho_{\hat{a}}(1)$ and the second, $\rho_{\hat{a}}(2)$.

Table 4.6*

Asymptotic values of $\rho_{\hat{a}}(1)$ and $\rho_{\hat{a}}(2)$

φ \ θ	-.95	-.50	.50	.95
-.95	-	.265 .085	-.381 -.03	-.481 -.036
-.50	.049 -.222	-	-.223 -.201	-.321 -.267
.50	.321 -.267	.223 -.201	-	-.049 -.222
.95	.481 -.036	.381 -.03	-.265 .085	-

We see that if the true value of φ is large in magnitude, residual autocorrelations would lead to the choice of an MA(1) model for $\hat{a}_t$ and therefore the correct identification. For intermediate values of φ, a moving average of order 2 or higher might be selected resulting in overparametrization.

In a recent paper [12], procedures have been proposed to determine the order of univariate ARMA models using functions of the autocorrelations of z_t other than those discussed above. The vector situation is a subject for further study.

---

*We are grateful to R.S. Tsay for computing this table.

4.2 Estimation

Once the order of the model in (3.1) has been tentatively selected, efficient estimates of the associated parameter matrices $\underset{\sim}{\varphi} = (\underset{\sim}{\varphi}_1,\dots,\underset{\sim}{\varphi}_p)$, $\underset{\sim}{\theta} = (\underset{\sim}{\theta}_1,\dots,\underset{\sim}{\theta}_q)$ and Σ are then determined by maximizing the likelihood function. Approximate standard errors and correlation matrix of the estimates of elements of the $\underset{\sim}{\varphi}$'s and $\underset{\sim}{\theta}$'s can also be obtained.

Conditional likelihood

For the ARMA (p,q) model, we can write

$$\underset{\sim}{a}_t = \underset{\sim}{z}_t - \underset{\sim}{\varphi}_1\underset{\sim}{z}_{t-1} - \dots - \underset{\sim}{\varphi}_p\underset{\sim}{z}_{t-p} + \underset{\sim}{\theta}_1\underset{\sim}{a}_{t-1} + \dots + \underset{\sim}{\theta}_q\underset{\sim}{a}_{t-q} \,. \tag{4.13}$$

As in the univariate case [7], the likelihood function can be approximated by a "conditional" likelihood function as follows. The series is regarded as consisting of the n-p vector observations $\underset{\sim}{z}_{p+1},\dots,\underset{\sim}{z}_n$. The likelihood function is then determined from $\underset{\sim}{a}_{p+1},\dots,\underset{\sim}{a}_n$ using the preliminary values $\underset{\sim}{z}_1,\dots,\underset{\sim}{z}_p$ and conditional on zero values for $\underset{\sim}{a}_p,\dots,\underset{\sim}{a}_{p-q-1}$. Thus,

$$\ell_c(\underset{\sim}{\varphi},\underset{\sim}{\theta},\Sigma|z) \propto |\Sigma|^{-\frac{n-p}{2}} \exp(-\tfrac{1}{2}\mathrm{tr}\,\Sigma^{-1} S(\underset{\sim}{\varphi},\underset{\sim}{\theta})) \tag{4.14}$$

where

$$S(\underset{\sim}{\varphi},\underset{\sim}{\theta}) = \sum_{t=p+1}^{n} \underset{\sim}{a}_t\underset{\sim}{a}_t' \,.$$

It has been shown in [14] that this approximation can be seriously inadequate if n is not sufficiently large and one or more zeros of $|\underset{\sim}{\theta}_q(B)|$ lie on or close to the unit circle. Specifically, this would lead to estimates of the moving average parameters with large bias.

This can be particularly troublesome with a seasonal model of the form

$$\underset{\sim}{\varphi}_p(B)\underset{\sim}{z}_t = \underset{\sim}{\theta}_q(B^s)\underset{\sim}{a}_t \tag{4.15}$$

for then the series can be approximately regarded as s separate component series each of length n/s. For example, for the simple univariate model $z_t = (1-\theta B)a_t$ with $\sigma^2 = 1$, the extent of the bias has been studied in [14] by considering the expected log likelihood

$$\underset{z|\theta_0}{E} \log\ell(\theta|\underset{\sim}{z}) = f_n(\theta,\theta_0) \tag{4.16}$$

where θ_0 is the true value. Values of θ which maximize $f_n(\theta,1)$ for various values of n are as follows:

n	10	50	100	1000
maximizing value of θ	.72	.87	.90	.97

Exact likelihood function

Exact likelihood function for the stationary vector ARMA(p,q) model has been derived in [14]. It takes the form

$$\ell(\underset{\sim}{\varphi},\underset{\sim}{\theta},\underset{\sim}{\Sigma}|\underset{\sim}{z}) = \ell_c(\underset{\sim}{\varphi},\underset{\sim}{\theta},\underset{\sim}{\Sigma}|\underset{\sim}{z})\ell_1(\underset{\sim}{\varphi},\underset{\sim}{\theta},\underset{\sim}{\Sigma}|\underset{\sim}{z}) \tag{4.17}$$

where ℓ_1 depends (i) only on $\underset{\sim}{z}_1,\ldots,\underset{\sim}{z}_p$ if q = 0 and (ii) on all the data vectors $\underset{\sim}{z}_1,\ldots,\underset{\sim}{z}_n$. Estimation algorithms have been developed and incorporated in our computer package for the vector MA(q) model where $\underset{\sim}{\theta}_q(B)$ assumes the multiplicative seasonal form

$$\underset{\sim}{\theta}_q(B) = \underset{\sim}{\theta}_{q_1}(B)(\underset{\sim}{I}-\underset{\sim}{\theta}_s B^s) \tag{4.18}$$

For the general ARMA(p,q) model, it has been shown that a close approximation to the exact likelihood can be obtained by considering the transformation

$$\underset{\sim}{w}_t = (\underset{\sim}{I}-\underset{\sim}{\varphi}_1 B-\ldots-\underset{\sim}{\varphi}_p B^p)\underset{\sim}{z}_t$$

so that (4.19)

$$\underset{\sim}{w}_t = \underset{\sim}{\theta}_q(B)\underset{\sim}{a}_t$$

and then apply the results for MA(q) to $\underset{\sim}{w}_t$, t = p+1,...,n.

In our computer package, both conditional likelihood (4.14) and exact likelihood via the approximation (4.19) are made available for parameter estimation. While estimates of moving average parameters using exact likelihood are superior, the computing time for the exact likelihood is, however, usually several times larger than that required for the conditional likelihood. We presently employ the conditional method in the preliminary stages of iterative model building and switch to the exact method towards the end.

4.3 Diagnostic checking

An important phase in the iterative model building process is model criticism. To guard against model misspecification and to search for directions of improvement, a detailed diagnostic analysis of the residual series $\{\hat{\underset{\sim}{a}}_t\}$ where

$$\hat{\underset{\sim}{a}}_t = \underset{\sim}{z}_t - \hat{\underset{\sim}{\varphi}}_1\underset{\sim}{z}_{t-1} - \ldots - \hat{\underset{\sim}{\varphi}}_p\underset{\sim}{z}_{t-p} + \hat{\underset{\sim}{\theta}}_1\hat{\underset{\sim}{a}}_{t-1} + \ldots + \hat{\underset{\sim}{\theta}}_q\hat{\underset{\sim}{a}}_{t-q} \tag{4.20}$$

should be performed. Our present computer package includes (a) plots of standardized residual series against time and (b) cross correlation matrices of the residuals $\hat{\underset{\sim}{a}}_t$. Again, the structure of the correlations

<u>Table 4.7</u>

Estimation and Diagnostic Checking for the Data in Figure 3.1 Corresponding to the MA(1) Model (4.21)

(a) Estimation results

<u>Conditional likelihood estimates</u> <u>Exact likelihood estimates</u>

(standard errors of estimates given in parenthesis)

$$\underset{\sim}{\theta}_0 = \begin{bmatrix} 17.12 \\ (.10) \\ 25.08 \\ (.08) \end{bmatrix} \qquad \begin{bmatrix} 17.07 \\ (.10) \\ 25.04 \\ (.08) \end{bmatrix}$$

$$\underset{\sim}{\theta} = \begin{bmatrix} .19 & .41 \\ (.04) & (.09) \\ -.57 & 1.12 \\ (.03) & (.07) \end{bmatrix} \qquad \begin{bmatrix} .18 & .43 \\ (.04) & (.10) \\ -.00 & 1.20 \\ (.03) & (.07) \end{bmatrix}$$

$$\underset{\sim}{\Sigma} = \begin{bmatrix} 4.87 & \\ 1.11 & 1.09 \end{bmatrix} \qquad \begin{bmatrix} 4.87 & \\ 1.13 & 1.01 \end{bmatrix}$$

(b) Pattern of cross correlation of residuals after exact likelihood fit

	a_1	a_2
a_1	....	.......-....
a_2	....	....

are summarized by assigning a plus sign to values greater than $2n^{-1/2}$, a minus sign for values less than $-2n^{-1/2}$ and a dot for in between cases.

As an illustration for the tools used in estimation and diagnostic checking, Table 4.7 gives the results for the series in Figure 3.1 using the tentatively specified MA(1) model

$$\underset{\sim}{Z}_t = \underset{\sim}{\theta}_0 + (\underset{\sim}{I} - \underset{\sim}{\theta} B)\underset{\sim}{a}_t \qquad (4.21)$$

where $\underset{\sim}{\theta}_0$ is the mean vector. Since the sample size is rather large, n = 250, there is scarcely any difference betwen the conditional likelihood estimates and the exact likelihood estimates. Also, the pattern of the symbols for the residual cross correlation indicates that the MA(1) model is adequate.

5. <u>Analyses of Three Actual Examples</u>

We now apply the *model building approach introduced in the preceding* section to three actual data sets

(i) The Financial Time Ordinary Share Index, U.K. Car Production and the Financial Time Commodity Price Index: Quarterly Data 3/1952-4/1967, obtained from [11]. This will be referred to as the SCC data.

(ii) The Gas Furnace Data given in [7].

(iii) The monthly Census Housing Data analyzed earlier in [14].

5.1 The SCC data

The three series are shown in Figure 5.1 where

Z_{1t}: Financial Time Ordinary Share Index

Z_{2t}: U.K. Car Production

Z_{3t}: Financial Time Commodity Price Index

In [11], the authors were interested in the possibility of predicting Z_{1t} from lagged values of Z_{2t} and Z_{3t} using a standard regression analysis in which Z_{1t} was treated as a dependent variables and $Z_{2(t-6)}$ and $Z_{3(t-7)}$ as regressors or independent variables. *For a discussion of this approach*, see [8].

We now consider what structure is revealed by the present multiple time series analysis, in which the three series are jointly modelled.

Tentative specification

A condensed summary of the pattern of cross correlations for the first 20 lags is provided in Table 5.1 in terms of the plus, minus and dot symbols. The original series show high and persistent auto and cross correlations. The standardized generalized partial cross correlations and related statistics up to the 5^{th} lag are given in Table 5.2. It is very clear from this analysis that little improvement occurs after lag $\ell = 2$. For $\ell > 2$, most of the elements of $\hat{\underset{\sim}{P}}(\ell)$ are small compared with their estimated standard errors, and the *$M(\ell)$ statistic, which is approximately* distributed as χ^2 with 9 degrees of freedom, fails to show significant improvement. *Table 5.3 gives the patterns of the cross correlations of*

Figure 5.1 The SCC Data

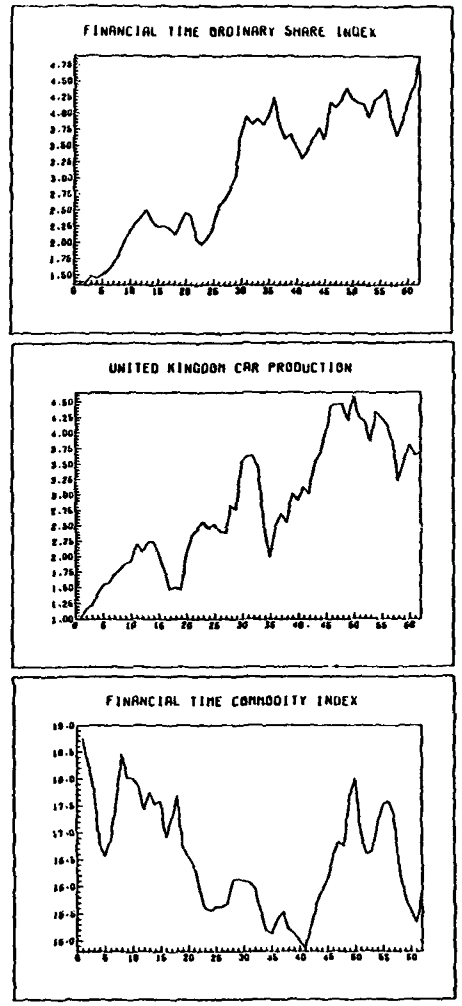


Table 5.1

Pattern of Sample Cross Correlations for the SCC Data

	Z_1 Stocks	Z_2 Cars	Z_3 Commodities
Z_1 Stocks	++++++++....	------------- --------	------------- --------
Z_2 Cars	....+++	+++++++++++++ +++.....	+++++++++++++ ++.....
Z_3 Commodities	-------..... ...+++++	+++++++++++++ ++++++..	+++++++++++++ +++++...

the residuals after AR(1) and AR(2) fit. The pattern after AR(2) is such as would be expected from "estimates white noise". Thus, an AR(2) model is suggested for the data. Alternately, note that there is one large residual correlation at lag 1 after an AR(1) fit, suggesting also the possibility of an ARMA(1,1) model.

Estimation

Both an AR(2) and an ARMA(1,1) model were fitted using the exact likelihood method (*although for this example, estimates from the conditional* likelihood for the ARMA(1,1) case are very close to the exact results). The ARMA(1,1) model produced a marginally better representation, therefore only its results are given. Specifically we use the model

$$(\underset{\sim}{I}-\underset{\sim}{\Phi}B)\underset{\sim}{Z}_t = \underset{\sim}{\theta}_0 + (\underset{\sim}{I}-\underset{\sim}{\Theta}B)\underset{\sim}{a}_t \qquad (5.1)$$

where $\underset{\sim}{\theta}_0$ is a vector of constants. Table 5.4 shows the initial unrestricted fit and also the fits after setting to zero those coefficients whose estimates are small compared to their standard errors.

Table 5.2

Standardized Partial Cross Correlations and Related Statistics: SCC Data

Lag	Standardized coefficients			Summary symbols	M(ℓ) Statistic χ^2_9	Diagonal elements of $\underset{\sim}{\Sigma}$ × 10
1	13.5 .4 - .8	1.3 10.5 .6	-1.8 -1.0 14.4	+ . . . + . . . +	301.3	.44 .89 1.62
2	-2.0 -1.1 -3.2	1.9 .1 2.2	- .1 -1.0 2.2	- . . . . . - + -	18.6	.40 .84 1.23
3	- .9 1.0 .5	2.0 -1.3 .2	-1.8 - .2 .6	. . . .	9.6	.37 .81 1.21
4	.2 .6 - .3	.3 .4 -.7	.5 - .9 .6	. . . .	3.6	.36 .79 1.19
5	- .8 .3 1.1	2.7 2.2 -.5	- .4 -1.6 1.6	. + . . + ...	11.9	.32 .70 1.11

Table 5.3

Pattern of Cross Correlation Matrices of Residuals: SCC Data

Lag	1	2	3	4	5	6	7	8
(a) AR(1) model								
	..+	... ...	... ...	... ...	... ...	... ...	... .+	... ...
(b) AR(2) model								
	... ...	... ...	... ...	... -.. ...	... ...	... ...	... .+	... ...

Table 5.4

Estimation Results for the Model (5.1): SCC Data
(Exact likelihood)

$$\begin{array}{cccc} \hat{\theta}_0 & \hat{\phi} & \hat{\theta} & \Sigma \end{array}$$

(1) Full Model

$$\begin{bmatrix} 1.11 \\ (.64) \\ 1.74 \\ (.82) \\ 4.08 \\ (1.47) \end{bmatrix} \quad \begin{bmatrix} .81 & .15 & -.06 \\ (.08) & (.07) & (.04) \\ -.07 & .98 & -.09 \\ (.10) & (.10) & (.05) \\ -.32 & .30 & .76 \\ (.18) & (.17) & (.08) \end{bmatrix} \quad \begin{bmatrix} -.29 & .23 & .06 \\ (.15) & (.11) & (.07) \\ -.45 & .20 & -.15 \\ (.22) & (.17) & (.11) \\ -.79 & .57 & -.44 \\ (.28) & (.21) & (.13) \end{bmatrix} \quad \begin{bmatrix} .037 & & \\ .022 & .078 & \\ .013 & .022 & .129 \end{bmatrix}$$

(2) Restricted Model (intermediate)

$$\begin{bmatrix} .13 \\ (.09) \\ .59 \\ (.05) \\ 2.48 \\ (1.10) \end{bmatrix} \quad \begin{bmatrix} .90 & .08 & \cdot \\ (.06) & (.06) & \\ \cdot & .92 & -.02 \\ & (.04) & (.04) \\ \cdot & \cdot & .85 \\ & & (.07) \end{bmatrix} \quad \begin{bmatrix} -.22 & .15 & \cdot \\ (.14) & (.10) & \\ -.31 & \cdot & \cdot \\ (.17) & & \\ -.55 & .22 & -.44 \\ (.23) & (.15) & (.12) \end{bmatrix} \quad \begin{bmatrix} .042 & & \\ .022 & .079 & \\ .017 & .021 & .131 \end{bmatrix}$$

(3) Restricted Model (final)

$$\begin{bmatrix} .12 \\ (.08) \\ .24 \\ (.10) \\ 2.76 \\ (1.07) \end{bmatrix} \quad \begin{bmatrix} .98 & & \cdot \\ (.03) & & \\ \cdot & .93 & \cdot \\ & (.04) & \\ \cdot & \cdot & .83 \\ & & (.06) \end{bmatrix} \quad \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ -.40 & \cdot & -.41 \\ (.23) & & (.12) \end{bmatrix} \quad \begin{bmatrix} .045 & & \\ .024 & .085 & \\ .019 & .023 & .134 \end{bmatrix}$$

Table 5.5

Pattern of Residual Cross Correlations After Final Restricted ARMA(1,1) Model Fit: SCC Data

	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$
$\hat{a}_1$	........-.	....	....
$\hat{a}_2$	....	....	....
$\hat{a}_3$	....	........-.	....

Diagnostic checking

Table 5.5 shows the pattern of residual cross correlations after the final restricted ARMA(1,1) fit. It suggests that the restricted model provide an adequate representation of the data.

Implication of the model

The final fitted results imply that the system is approximately represented by

$$(1-.98B)Z_{1t} = a_{1t} \tag{5.2a}$$

$$(1-.93B)Z_{2t} = .2 + a_{2t} \tag{5.2b}$$

$$(1-.83B)Z_{3t} = 2.8 + .40a_{1(t-1)} + (1+.41B)a_{3t} \ . \tag{5.2c}$$

Upon substituting (5.2a) in (5.2c), we get

$$(1-.83B)Z_{3t} = 2.8 + .40(1-.98B)Z_{1(t-1)} + (1+.41B)a_{3t} \ . \tag{5.2d}$$

Thus all three series behave approximately as random walks with slightly correlated innovations. From the point of view of forecasting, (5.2d) is of some interest since it implies that ordinary share $Z_{1(t-1)}$ is a leading indicator at lag 1 for the commodity index Z_{3t}. Its effect, however, is small as can be seen for example by the improvement achieved over the corresponding best fitting univariate model which was

$$(1-.78B)Z_{3t} = 3.63 + (1+.53B)a_t, \ \sigma^2 = .151 \tag{5.3}$$

The residual variance of 0.151 from the univariate model is not much larger than the value 0.134 for a_{3t} obtained from the final vector model. Although the multiple time series analysis fails to reveal anything very surprising for this example, it shows what is there and does not mislead.

5.2 The Gas Furnace Data

The two series shown in Figure 5.2 consist of (i) input gas rate and (ii) output as CO_2 concentration at 9 second intervals from a gas furnace. We shall let Z_{1t} = gas rate + .057 and $Z_{2t} = CO_2 - 5.35$. This set of data was employed in [7] to illustrate a transfer function modelling procedure. The procedure was designed to model the dynamic relationship of two stochastic time series *one of which is* <u>*known*</u> *to be input for the other*. According to this approach, the appropriate model for the input Z_{1t} and that for the output Z_{2t} are, respectively

$$(1-1.97B+1.37B^2-.34B^3)Z_{1t} = \alpha_t\,, \qquad \hat{\sigma}_\alpha^2 = .0353 \tag{5.4a}$$

and

$$Z_{2t} = \nu(B)Z_{1t} + (1-1.53B+.63B^2)^{-1}a_t\,, \quad \hat{\sigma}_a^2 = .0561 \tag{5.4b}$$

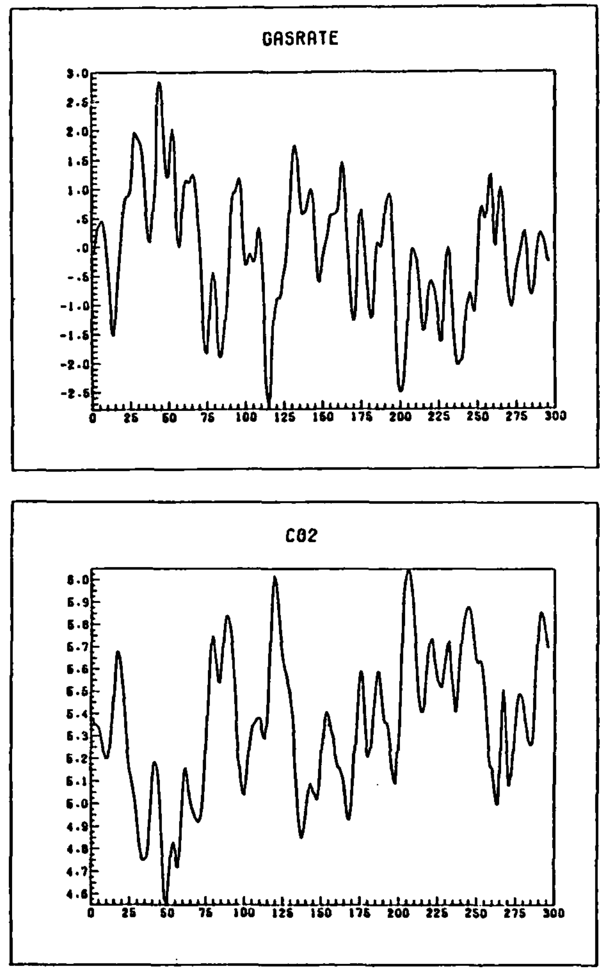
where $\nu(B)$ is the transfer function

$$\nu(B) = \frac{-(.53+.37B+.51B^2)}{1-.57B}\,B^b, \tag{5.4c}$$

$b = 3$ is the delay and the $\{a_t\}$ and $\{\alpha_t\}$ series are assumed independent.

It will be of interest to analyze the data using the present approach where no distinction is made between an input and an output variable.

Figure 5.2 The Gas Furnace Data



Tentative specification

In Table 5.6, part (a) shows that the auto and cross correlations of the original data are persistently large in magnitude as the lag increases ruling out low order vector moving average models; part (b) gives the $M(\ell)$ statistic, which should be compared with a χ^2 variable with 4 degrees of freedom, through the 11th lag, suggesting that an AR(6) model might be tentatively selected; and part (c) shows the residual cross correlation pattern after an AR(6) fit, confirming the appropriateness of this model.

Table 5.6

Tentative Identification for the Gas Furnace Data

(a) *Pattern of cross correlations of the original data*

	Z_{1t}	Z_{2t}
Z_{1t}	++++++++++++	------------
Z_{2t}	----------..	++++++++++++

(b) M statistic for generalized partial cross correlations

Lag ℓ	1	2	3	4	5	6	7	8	9	10	11
$M(\ell)$	1650	665	31.7	22.5	5.6	12.9	1.8	8.0	3.5	0	2.0

(c) Pattern of cross correlations of the residuals after AR(6) fit

	$\hat{a}_{1t}$	$\hat{a}_{2t}$
$\hat{a}_{1t}$	............-	....
$\hat{a}_{2t}$	....	....

Estimation results

Estimation results corresponding to an unrestricted AR(6) model

$$(\underline{I}-\underline{\varphi}_1 B-\dots-\underline{\varphi}_6 B^6)\underline{Z}_t = \underline{a}_t \tag{5.5}$$

are as follows:

$$\begin{array}{cccccc} \hat{\underline{\varphi}}_1 & \hat{\underline{\varphi}}_2 & \hat{\underline{\varphi}}_3 & \hat{\underline{\varphi}}_4 & \hat{\underline{\varphi}}_5 & \hat{\underline{\varphi}}_6 \\ \begin{bmatrix} 1.93 & -.05 \\ (.06) & (.05) \\ .06 & 1.55 \\ (.08) & (.06) \end{bmatrix} & \begin{bmatrix} -1.20 & .10 \\ (.13) & (.08) \\ -.14 & -.59 \\ (.16) & (.11) \end{bmatrix} & \begin{bmatrix} .17 & -.08 \\ (.15) & (.09) \\ -.44 & -.17 \\ (.19) & (.11) \end{bmatrix} & \begin{bmatrix} -.16 & .03 \\ (.15) & (.09) \\ .15 & .13 \\ (.19) & (.11) \end{bmatrix} & \begin{bmatrix} .38 & -.04 \\ (.14) & (.08) \\ -.12 & .06 \\ (.18) & (.10) \end{bmatrix} & \begin{bmatrix} -.22 & .03 \\ (.08) & (.03) \\ .25 & -.04 \\ (.11) & (.04) \end{bmatrix} \end{array}$$

$$\underline{\Sigma} = \begin{bmatrix} .0345 & \\ -.0023 & .0566 \end{bmatrix}, \quad \hat{\rho}(a_1,a_2) = .045 \tag{5.6}$$

If we let

$$\hat{\underline{\varphi}}_\ell = (\hat{\varphi}_{ij}^{(\ell)})$$

then we see that $\hat{\varphi}_{12}^{(\ell)}$ are small compared with their standard errors over all lags, confirming (as in this case is known from the physical nature of the apparatus generating the data) that there is a unidirectional relationship between Z_{1t} and Z_{2t} involving no feedback. Also, $\hat{\varphi}_{21}^{(\ell)}$ is small for $\ell = 1,2$, and the residuals $\hat{a}_{1t}$ and $\hat{a}_{2t}$ are essentially uncorrelated, implying a delay of 3 periods. It should be noted in addition, that the variances for a_{1t} and a_{2t} are very close to those for α_t and a_t in (5.4).

To facilitate comparison with the previous results in (5.4), we set $\varphi_{11}^{(\ell)} = 0$ for $\ell > 3$, $\varphi_{12}^{(\ell)} = 0$ for all ℓ, $\varphi_{21}^{(\ell)} = 0$ for $\ell = 1,2$ and

$\varphi_{22}^{(\ell)} = 0$ for $\ell = 5,6$. Estimation results for this restricted AR(6) model are then

$$\begin{array}{cccccc} \hat{\varphi}_1 & \hat{\varphi}_2 & \hat{\varphi}_3 & \hat{\varphi}_4 & \hat{\varphi}_5 & \hat{\varphi}_6 \\ \begin{bmatrix} \underset{(.06)}{1.98} & \cdot \\ \cdot & \underset{(.06)}{1.53} \end{bmatrix} & \begin{bmatrix} \underset{(.10)}{-1.38} & \cdot \\ \cdot & \underset{(.11)}{-.58} \end{bmatrix} & \begin{bmatrix} \underset{(.06)}{.35} & \cdot \\ \underset{(.07)}{-.53} & \underset{(.10)}{-.14} \end{bmatrix} & \begin{bmatrix} \cdot & \cdot \\ \underset{(.16)}{.11} & \underset{(.04)}{.12} \end{bmatrix} & \begin{bmatrix} \cdot & \cdot \\ \underset{(.17)}{-.04} & \cdot \end{bmatrix} & \begin{bmatrix} \cdot & \cdot \\ \underset{(.11)}{.21} & \cdot \end{bmatrix} \end{array}$$

$$\hat{\Sigma} = \begin{bmatrix} .0359 & -.0029 \\ & .0561 \end{bmatrix}, \quad \hat{\rho}(a_1,a_2) \doteq 0 \tag{5.7}$$

Examination of the pattern of the cross correlations of the residuals suggest that the model is adequate.

Implication of the bivariate model

From (5.7), the equation for the input is

$$(1-1.98B+1.38B^2-.35B^3)Z_{1t} = a_{1t} \tag{5.8}$$

with $Var(a_{1t}) = .0359$, which is essentially the same as (5.4a). The output model is

$$Z_{2t} = \nu_*(B) + (1-1.53B+.58B^2+.14B^3-.12B^4)^{-1}a_{2t} \tag{5.9}$$

with $Var(a_{2t}) = .0561$, where the transfer function $\nu_*(B)$ is

$$\nu_*(B) = \frac{-.53+.11B+.21B^3}{(1-1.53B+.58B+.14B^3-.12B^4)} B^3 \tag{5.10}$$

The noise model which is the second term on the right hand side of (5.9) is not very different from the corresponding one in (5.4b). Except for

the delay, the form of the transfer function $\nu_*(B)$ appears markedly different from that of $\nu(B)$ in (5.4c). However, if we expand these two ratios of polynomials, $\nu(B) = \sum_{j=0}^{\infty} \nu_j B^j$ and $\nu_*(B) = \sum_{j=0}^{\infty} \nu_j^* B^j$ the impulse response weights ν_j's and ν_j^*'s are in fact quite similar as shown in the following table.

Table 5.7

Impulse Response Weights for the Gas Furnace Data

j	0	1	2	3	4	5	6	7	8	9	10	11	12
ν_j	.	.	.	-.53	-.67	-.89	-.51	-.29	-.17	-.09	-.05	-.03	-.02
ν_j^*	.	.	.	-.53	-.70	-.77	-.48	-.26	-.09	-.01	.01	.00	-.01

Thus, the bivariate model (5.7) and the results in (5.4) are in essential agreement.

Further analysis of stepwise AR results

As mentioned earlier, in estimating the generalized partial cross correlations, autoregressive models of increasing order are successively fitted by least squares. For the gas furnace data, it is of interest to examine the changes in the fitting results as the order is increased. Table 5.8 shows the situation for $p = 1,\ldots,6$. Instead of giving the estimates, a plus (minus)sign is used when the estimate is greater (less) than 2 (-2) times its standard error and a blank for in between values. The residual covariance matrix for each order is also given.

Table 5.8

Order of AR	$\underset{\sim}{\varphi}_1$	$\underset{\sim}{\varphi}_2$	$\underset{\sim}{\varphi}_3$	$\underset{\sim}{\varphi}_4$	$\underset{\sim}{\varphi}_5$	$\underset{\sim}{\varphi}_6$	$\underset{\sim}{\Sigma}$
1	$\begin{bmatrix} + & + \\ - & + \end{bmatrix}$						$\begin{bmatrix} .102 & .090 \\ & .346 \end{bmatrix}$
2	$\begin{bmatrix} + & - \\ + & + \end{bmatrix}$	$\begin{bmatrix} - & + \\ - & - \end{bmatrix}$					$\begin{bmatrix} .037 & -.004 \\ & .069 \end{bmatrix}$
3	$\begin{bmatrix} + & \\ & + \end{bmatrix}$	$\begin{bmatrix} - & \\ - & - \end{bmatrix}$	$\begin{bmatrix} + & \\ & + \end{bmatrix}$				$\begin{bmatrix} .036 & -.002 \\ & .063 \end{bmatrix}$
4	$\begin{bmatrix} + & \\ & + \end{bmatrix}$	$\begin{bmatrix} - & \\ & - \end{bmatrix}$	$\begin{bmatrix} & \\ - & \end{bmatrix}$	$\begin{bmatrix} & \\ + & + \end{bmatrix}$			$\begin{bmatrix} .036 & -.003 \\ & .059 \end{bmatrix}$
5	$\begin{bmatrix} + & \\ & + \end{bmatrix}$	$\begin{bmatrix} - & \\ & - \end{bmatrix}$	$\begin{bmatrix} & \\ - & \end{bmatrix}$	$\begin{bmatrix} & \\ & \end{bmatrix}$	$\begin{bmatrix} & \\ & \end{bmatrix}$		$\begin{bmatrix} .035 & -.003 \\ & .058 \end{bmatrix}$
6	$\begin{bmatrix} + & \\ & + \end{bmatrix}$	$\begin{bmatrix} - & \\ & - \end{bmatrix}$	$\begin{bmatrix} & \\ - & \end{bmatrix}$	$\begin{bmatrix} & \\ & \end{bmatrix}$	$\begin{bmatrix} + & \\ & \end{bmatrix}$	$\begin{bmatrix} - & \\ + & \end{bmatrix}$	$\begin{bmatrix} .035 & -.002 \\ & .057 \end{bmatrix}$

The following observations may be made.

(i) If only AR(1) or AR(2) were contemplated, one might be lead to believe mistakenly that there was a feedback relationship between these two series.

(ii) The dynamic transfer function relationship becomes clear when the order of the model, p, is increased to three. Since the input series Z_{1t} essentially follows a univariate AR(3) model, this suggests that the present procedure would correctly identify the one-sided causal dynamic relationship after the input model were appropriately selected.

(iii) The delay b = 3 emerges when the order p is increased to 4. Since only very marginal improvement in the fit occurs for p > 4, this is saying that the delay is correctly identified when the model is essentially correctly specified.

Implications on general transfer function model building

While the transfer function modelling procedure proposed in [7] has been found useful for one output and one input, it becomes rather complex when we have more than one input variable. The Gas Furnace example discussed above suggests that the present multiple time series procedure may provide a useful alternative. It seems to possess the following advantages.

(i) Specification of a model seems more direct and straightforward. The one-sided causal relationship will emerge in the identification process, and the stochastic structures of the input as well as the transfer function relationship between input and output are modelled simultaneously.

(ii) More important, it can readily handle multiple input and multiple output situations, assuming that the inputs are stochastic. As mentioned earlier, writing $\underset{\sim}{Z}_t' = (\underset{\sim}{Z}_{1t}', \underset{\sim}{Z}_{2t}')$ a general transfer function relationship between the input vector $\underset{\sim}{Z}_{1t}$ and the output vector $\underset{\sim}{Z}_{2t}$ occurs as a special case of (3.1) when all the $\underset{\sim}{\varphi}$'s and $\underset{\sim}{\theta}$'s are lower block triangular.

(iii) The data are allowed to shed light on the existence of the transfer function relationship in the specification process. In economic and business applications, this would provide a useful way to search for leading indicators. Conversely, this one-sided dynamic relationship may not exist between two time series even when one variable is known to be the input for the other. One reason for this phenomenon is the effect of temporal aggregation.

For example, suppose that on a monthly basis the output, say, consumption y_t is related to the input income x_t by

$$y_t = \nu x_{t-1} + e_t \tag{5.11}$$

and that x_t follows the model

$$x_t = \phi x_{t-1} + \alpha_t$$

where $\{e_t\}$ and $\{\alpha_t\}$ are two white noise processes independent of each other. Now instead of monthly data suppose only quarterly totals of income and consumption are available. Writing $\underset{\sim}{Z}_T = (Z_{1T}, Z_{2T})'$ where Z_{1t} is the quarterly consumption and Z_{2T} the quarterly income, it is shown in [17] that $\underset{\sim}{Z}_T$ follows the ARMA(1,1) model where neither $\underset{\sim}{\varphi}$ nor $\underset{\sim}{\theta}$ is triangular. In other words, psuedo feedback relationships could occur because of this temporal aggregation effect, and it would be a mistake to impose a transfer function model in such a situation.

5.3 Census housing data

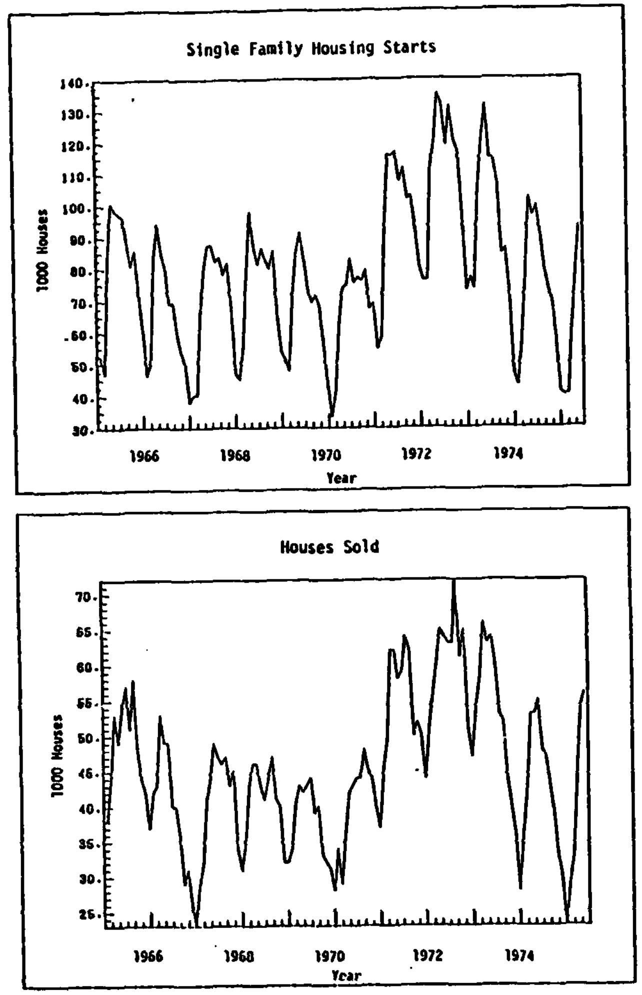
As a third example, we consider the monthly Housing Starts Z_{1t} and Houses Sold Z_{2t} shown in Figure 5.3. These two series were obtained from the Bureau of the Census and an earlier analysis was given in [14].

Tentative specification

Because of the strong seasonal behavior of the series, one may very well be led to consider the seasonally differenced series $\underset{\sim}{U}_t = (U_{1t}, U_{2t})'$ where

$$U_{1t} = (1-B^{12})Z_{1t}$$
$$U_{2t} = (1-B^{12})Z_{2t}$$

Figure 5.3 Census Housing Data, Monthly January 1965-May 1975



Part (a) of Table 5.9 shows the pattern of the cross correlations of U_t indicating that low order vector MA model would not be appropriate. Part (b) of the same table gives the $M(\ell)$ statistics for $\ell = 1,\ldots,5$, and Part (c) shows the pattern of the cross correlations of the residuals after an AR(1) fit. These summaries suggest the tentative model

$$(\underset{\sim}{I}-\underset{\sim}{\Phi}B)(1-B^{12})\underset{\sim}{Z}_t = (\underset{\sim}{I}-\underset{\sim}{\Theta}B^{12})\underset{\sim}{a}_t. \tag{5.12}$$

Table 5.9

Tentative Identification for the Seasonally Differenced Housing Data $\underset{\sim}{U}_t$

(a) Pattern of cross correlations

	U_{1t}	U_{2t}
U_{1t}	++++++++++.. ......------ ------------	++++++++.... ....--- ----------..
U_{2t}	++++++++++.. ....--- -----------.	++++++++.... ....---------..

(b) $M(\ell)$ statistics corresponding to partial cross correlation matrices

Lag ℓ	1	2	3	4	5
$M(\ell)$	218.6	3.5	2.3	4.7	5.4

(c) Pattern of residual cross correlations after AR(1) fit

	a_{1t}	a_{2t}
a_{1t}	...........-	....+.
a_{2t}	....+..+...... ..-..-......	...........- ....

Estimation and checking

Table 5.10 summaries the estimation results corresponding to

(i) the full model in (5.12) using the conditional likelihood method,

(ii) the full model using the exact likelihood method,

(iii) restricted model by setting "small" parameter estimates to zero.

Table 5.11 shows the pattern of the cross correlations of the residuals corresponding to the restricted case, showing that the model gives an adequate representation of the series.

Table 5.10

Estimation Results for the Model (5.12): Census Housing Data

	(1) Full model conditional likelihood	(2) Full model Exact likelihood	(3) Restricted model Exact likelihood
$\underset{\sim}{\Phi} =$	$\begin{bmatrix} .47 & .89 \\ (.07) & (.13) \\ .14 & .69 \\ (.05) & (.08) \end{bmatrix}$	$\begin{bmatrix} .46 & .95 \\ (.07) & (.14) \\ .10 & .76 \\ (.05) & (.09) \end{bmatrix}$	$\begin{bmatrix} .42 & 1.03 \\ (.07) & (.13) \\ \cdot & .93 \\ & (.04) \end{bmatrix}$
$\underset{\sim}{\Theta} =$	$\begin{bmatrix} .75 & .06 \\ (.07) & (.11) \\ .07 & .69 \\ (.05) & (.08) \end{bmatrix}$	$\begin{bmatrix} 1.01 & -.04 \\ (.07) & (.12) \\ .05 & .99 \\ (.05) & (.07) \end{bmatrix}$	$\begin{bmatrix} .94 & \cdot \\ (.06) & \\ \cdot & 1.00 \\ & (.06) \end{bmatrix}$
$\underset{\sim}{\Sigma} =$	$\begin{bmatrix} 37.51 & \\ 6.29 & 15.15 \end{bmatrix}$	$\begin{bmatrix} 28.09 & \\ 4.98 & 11.13 \end{bmatrix}$	$\begin{bmatrix} 29.75 & \\ 5.89 & 11.83 \end{bmatrix}$
			$\hat{\rho}(a_1,a_2) \doteq .31$

Table 5.11

Residual Cross Correlations for the Restricted Model: Census Housing Data

	a_{1t}	a_{2t}
a_{1t}	....	......+..... ....
a_{2t}	....+-.....	....+-...-

Interpretation

For the full model, comparing the results of the conditional likelihood with those from the exact likelihood, we see that

(i) there is a substantial increase in the estimated values of the diagonal elements of $\underset{\sim}{\theta}$ to near unity, and a corresponding decrease in the variances of the residuals, when the exact method is used,

(ii) in contrast, little change occurs in the estimates of elements of $\underset{\sim}{\phi}$.

This is in agreement with the discussion earlier in Section 4.

Now from the restricted model, we can write, for the Houses Sold series Z_{2t}

$$(1-.93B)(1-B^{12})Z_{2t} = (1-B^{12})a_{2t} \tag{5.13a}$$

or

$$(1-.93B)Z_{2t} = S_{2t} + a_{2t} \tag{5.13b}$$

where S_{2t} satisfies the relation $S_{2t} = S_{2(t-12)}$. Thus, Z_{2t} behaves nearly like a random walk with a deterministic seasonal component and does not depend on the past of Z_{1t}.

On the other hand, for the Housing Starts series Z_{1t}, we have that

$$(1-.42B)(1-B^{12})Z_{1t} = 1.03(1-B^{12})Z_{2(t-1)} + (1-.94B^{12})a_{1t} \tag{5.14a}$$

so that the seasonal differencing operator $(1-B^{12})$ again nearly cancels yielding approximately

$$(1-.42B)Z_{1t} = S_{1t} + Z_{2(t-1)} + a_{1t} \tag{5.14b}$$

where $S_{1t} = S_{1(t-12)}$. Thus, Housing Starts Z_{1t} depends not only on its

own past, $Z_{1(t-1)}$, but also on the past of Houses Sold, $Z_{2(t-1)}$. It was quoted in [14] that an appropriate individual model for Z_{1t} was

$$(1-B)(1-B^{12})Z_{1t} = (1-.28B)(1-.91B^{12})c_t \tag{5.15a}$$

or approximately

$$(1-B)Z_{1t} = S_{1t} + (1-.28B)c_t,\ \sigma_c^2 = 41.61 \tag{5.15b}$$

We see from (5.13b) and (5.14b) that the difference operator (1-B) in (5.15) indicating that Z_{1t} is nonstationary, arises because of the dependence of Z_{1t} on $Z_{2(t-1)}$. Also, by comparing the variance σ_c^2 in (5.15b) with the corresponding variance of a_{1t} in Table 4.10, we see a substantial reduction when the information $Z_{2(t-1)}$ is utilized.

In summary, this example shows that (i) the existence of a deterministic seasonal component can be detected when the exact likelihood method is employed and (ii) an appreciable reduction in the one step ahead forecast variance can occur by modelling several series jointly.

6. Eigenvalue-Eigenvector Analyses in Multiple Time Series

In this section, we describe several types of eigenvalue-eigenvector analyses which have been found useful in analyzing multiple time series. Writing (3.1) in the form

$$\underset{\sim}{z}_t = \hat{\underset{\sim}{z}}_{t-1}(1) + \underset{\sim}{a}_t \tag{6.1}$$

where $\hat{\underset{\sim}{z}}_{t-1}(1)$ is the one step ahead forecast of $\underset{\sim}{z}_t$ made at time t-1, and denoting, for stationary series,

$$\underset{\sim}{\Gamma}_z(0) = E(\underset{\sim}{z}_t\underset{\sim}{z}_t') \text{ and } \underset{\sim}{\Gamma}_{\hat{z}}(0) = E(\hat{\underset{\sim}{z}}_{t-1}(1)\hat{\underset{\sim}{z}}_{t-1}(1)'),$$

it will often be informative to compute eigenvalues and eigenvectors of estimates of the following matrices

$$\text{(a) } \underset{\sim}{\Gamma}_z(0), \text{ (b) } \Sigma, \text{ (c) } \underset{\sim}{\Gamma}_z(0)^{-1}\underset{\sim}{\Gamma}_{\hat{z}}(0), \text{ (d) } \underset{\sim}{\Phi}_\ell \text{ and } \underset{\sim}{\theta}_\ell$$

Such analyses have two principal aims:

(i) detection of exact linear relations between series,

(ii) to aid understanding and interpretation of the fitted model.

6.1 Exact contemporaneous linear relationship

Suppose there are m zero eigenvalues in $\underset{\sim}{\Gamma}_z(0)$. This implies that there are m independent exact linear relationships of the form

$$\underset{\sim}{c}_j'\underset{\sim}{z}_t = 0 \qquad (6.2)$$

where $\underset{\sim}{c}_j' = (c_{1j},\dots,c_{kj})$, existed between the elements of $\underset{\sim}{z}_t$, $t = 1,\dots,n$. Such relationships occur when one or more series is computed from contemporaneous values of the others. Ideally, the analyst should know his data sources sufficiently well that such relationships will be known in advance. However, experience shows that this check should always be made at the initial specification stage with multivariate data, [3]. In this way the form of such relationships are forced to the attention and also by limiting subsequent analysis to linearly independent series, estimation computations are not frustrated by singularities. Eigenvalues which are close to zero can also warn of approximate contemporaneous linear relationships in the data.

In our computer package, eigenvalues and eigenvectors of the estimates of $\underset{\sim}{\Gamma}_z(0)$ are given in the output of the Preliminary Analysis Program.

6.2 Exact lagged linear relationship

Zero eigenvalues in the covariance matrix of $\underset{\sim}{a}_t$, Σ, arise when there are linear relationships of the form

$$\underset{\sim}{h}_1'\underset{\sim}{z}_t + \underset{\sim}{h}_2'\underset{\sim}{z}_{t-1} + \dots + \underset{\sim}{h}_r'\underset{\sim}{z}_{t-r} = 0, \qquad (6.3)$$

where the $\underset{\sim}{h}$'s are k×1 vector of constants, existed between elements in the series which are not concurrent. Equivalently, they indicate that the k series $\{\underset{\sim}{z}_t\}$ are driven by less than k innovation series. It is possible, for instance, for two series that look quite different to be generated by identical innovations passing through different filters. In practice, Σ can be estimated initially by fitting an autoregressive model of suitably high order. Specifically, it can be readily shown that the relationship in (6.3) implies that the residual covariance matrix of an AR(p) model will only be singular for $p \geq r$.

Eigenvalues and eigenvectors of $\hat{\Sigma}$ are available after each AR fit in the Stepwise Autoregression Program of our package. We now illustrate this by an example. Figure 6.1 shows 4 monthly industrial series, from January 1958 to July 1978, obtained from the Bureau of the Census. They are: z_{1t} - new orders, z_{2t} - shipments, z_{3t} - inventory and z_{4t} - unfilled orders. The new orders series were actually constructed from the shipments and unfilled orders series as follows

$$z_{1t} = z_{2t} + z_{4t} - z_{4(t-1)} \qquad (6.4)$$

although initially this relationship was not revealed to us. For an AR(1) model

$$\underset{\sim}{z}_t = \underset{\sim}{\Phi}\underset{\sim}{z}_{t-1} + \underset{\sim}{a}_t \qquad (6.5)$$

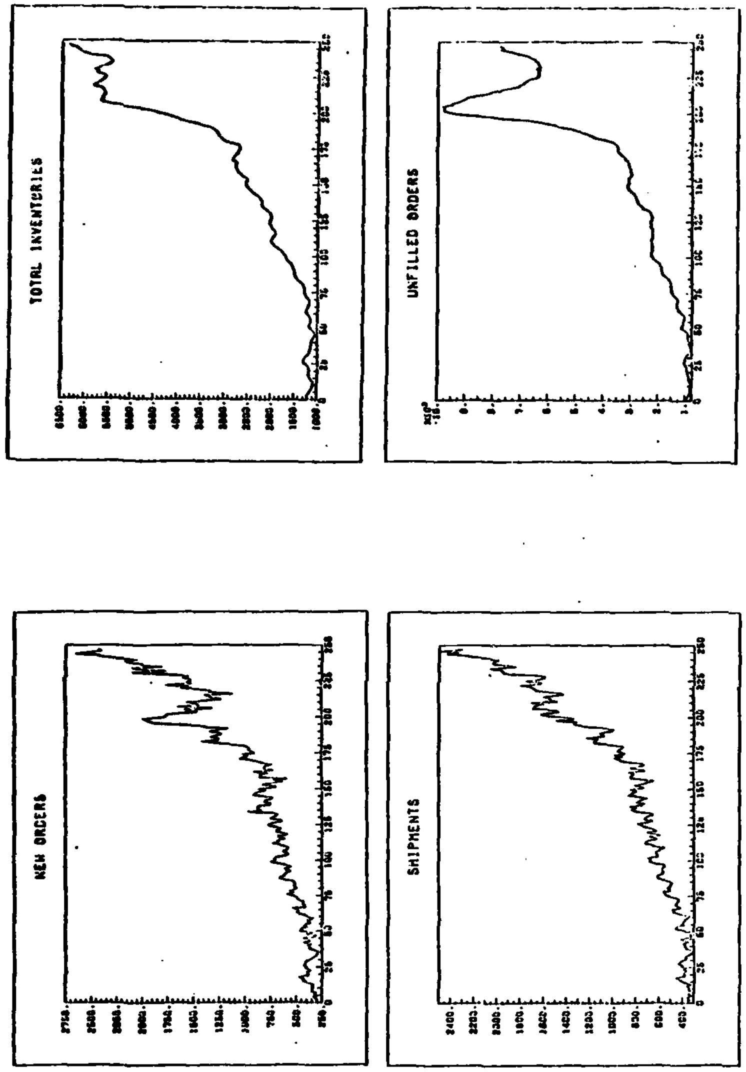


Figure 6.1 The Census Industry Series 1/58-7/78

the estimate of $\underset{\sim}{\varphi}$ is found to be

$$\hat{\underset{\sim}{\varphi}} = \begin{bmatrix} .23 & .75 & -.08 & -.02 \\ .14 & .51 & -.01 & .03 \\ -.05 & .03 & .95 & .04 \\ .09 & .24 & -.07 & .95 \end{bmatrix}$$

and the eigenvalues of $\hat{\underset{\sim}{\Sigma}}$ are $(52.1,\ 10.7,\ 5.2,\ -.5\times10^{-4})$. The eigenvector corresponding to the zero value ($-.5\times10^{-4}$ because of roundoff errors) is proportional to $\underset{\sim}{g}' = (-1, 1, 0, 1)$. Since $\underset{\sim}{g}'\underset{\sim}{a}_t \doteq 0$, we have that

$$\underset{\sim}{g}'\underset{\sim}{z}_t = \underset{\sim}{g}'\underset{\sim}{\varphi}\underset{\sim}{z}_{t-k} \quad . \tag{6.6}$$

Substituting $\hat{\underset{\sim}{\varphi}}$ for $\underset{\sim}{\varphi}$ we obtain almost exactly the relationship in (6.4).

6.3 A canonical analysis

Assuming stationarity, we have from (6.1) that

$$\underset{\sim}{\Gamma}_z(0) = \underset{\sim}{\Gamma}_{\hat{z}}(0) + \underset{\sim}{\Sigma} \tag{6.7}$$

In [10], a canonical analysis of the general model (6.1) with $\underset{\sim}{\Gamma}_z(0)$ and $\underset{\sim}{\Sigma}$ assumed positive definite leads to finding eigenvalues and eigenvectors of $\underset{\sim}{\Gamma}_z(0)^{-1}\underset{\sim}{\Gamma}_{\hat{z}}(0)$ (or equivalently of $\underset{\sim}{\Gamma}_z(0)^{-1}\underset{\sim}{\Sigma}$). Let $(\lambda_1,\ldots,\lambda_k)$ be the eigenvalues ordered with λ_1 the smallest and the $k\times k$ matrix $\underset{\sim}{M}' = (\underset{\sim}{m}_1,\ldots,\underset{\sim}{m}_k)$ consists of the corresponding eigenvectors. Then the transformed process $\underset{\sim}{y}_t = \underset{\sim}{M}\underset{\sim}{z}_t$ is such that

$$\underset{\sim}{y}_t = \hat{\underset{\sim}{y}}_{t-1}(1) + \underset{\sim}{b}_t$$

with

$$\underset{\sim}{\Gamma}_y(0) = \underset{\sim}{\Gamma}_{\hat{y}}(0) + \underset{\sim}{\Sigma}_b \tag{6.8}$$

where $\underset{\sim}{\Gamma}_y(0)$, $\underset{\sim}{\Gamma}_{\hat{y}}(0)$ and $\underset{\sim}{\Sigma}_b$ are all diagonal. Thus,

$$\lambda_i = \mathrm{Var}(\hat{y}_{i(t-1)}(1))/\mathrm{Var}(y_{it}) \quad i = 1,\ldots,k$$

is a measure of forecastability of the series. Suppose there are k_1 eigenvalues close to unity, k_3 eigenvalues close to zero, and k_2 eigenvalues intermediate in size. This implies (a) there is a near nonstationary space of dimension k_1 accounting for the overall growth factors in the series, and (b) that there are k_3 stable relationships among the variables which vary independently about fixed means.

An application of this analysis to a 5-variate U.S. hog data is given in [10]. It was found that the structure of the five series are greatly simplified by the transformation and that the transformed series are scientifically meaningful.

This canonical analysis is presently included in our Stepwise Autoregression Program. We plan to incorporate it in the Estimation and Forecasting Program in the near future.

6.4 Decoupling canonical transformation

Consider first the vector AR(1) process

$$\underset{\sim}{z}_t = \underset{\sim}{\varphi}\underset{\sim}{z}_{t-1} + \underset{\sim}{a}_t. \tag{6.9}$$

If $\underset{\sim}{\varphi}$ has k linearly independent eigenvectors, then there exists a k×k real matrix $\underset{\sim}{Q}$ such that $\underset{\sim}{Q}\underset{\sim}{\varphi}\underset{\sim}{Q}^{-1} = \underset{\sim}{\Lambda}$ where $\underset{\sim}{\Lambda}$ is a block diagonal matrix with the block size being 1 corresponding to real eigenvalues or 2 for complex pair of eigenvalues. Writing $\underset{\sim}{y}_t = \underset{\sim}{Q}\underset{\sim}{z}_t$ and $\underset{\sim}{b}_t = \underset{\sim}{Q}\underset{\sim}{a}_t$ we obtain

$$\underset{\sim}{y}_t = \underset{\sim}{\Lambda}\underset{\sim}{y}_{t-1} + \underset{\sim}{b}_t \tag{6.10}$$

When an eigenvalue of $\underset{\sim}{\varphi}$ is real, the corresponding transformed series is uncoupled in the sense that it may be optimally forecast from its own past. The existence of a complex pair of eigenvalues implies that only paired uncoupling is possible. That is that a pair of transformed series can be forecast by the past of only that pair. This procedure can be similarly applied to a MA(1) process. It can also be extended to cover the AR(p) model by first writing the k dimensional process as the AR(1) form for the k·p×1 dimensional vector $(\underset{\sim}{z}_t', \ldots, \underset{\sim}{z}_{t-p+1}')'$. Properties of such an extension are being investigated.

7. Alternative Approaches to Modelling Multiple Time Series

In Sections 4 and 5, we have proposed an approach to modelling multiple time series illustrated by several examples. We shall conclude this report by briefly discussing some other approaches which have been proposed and used in practice.

7.1 Cross correlating prewhitened residuals

This approach, originally proposed in [5], consists of first building an individual time series model for each series, and then attempting to identify the relationships between the series via studying the dynamic structure of the individually whitened residuals. Specifically, for k series $\{z_{1t}\}, \ldots, \{z_{kt}\}$, one first builds univariate models of the ARMA form

$$\varphi_{p_j}(B)z_{jt} = \theta_{q_j}(B)c_{jt} \quad j = 1,\ldots,k \tag{7.1}$$

and then cross correlates the k residual series $\{c_{1t}\},\ldots,\{c_{kt}\}$ to determine their dynamic structure from which the relationships among the $\{z_{jt}\}$ series are then deduced. This approach seems less satisfactory compared with the present one for two reasons:

(i) Complexity in relationships between prewhitened residuals: It can be shown that even if the vector series $\{z_t\}$ follows a low order ARMA model (3.1), the corresponding model for the vector series $\{c_t\}$, where $c_t = (c_{1t},\ldots,c_{kt})'$ can be rather complex and difficult to identify in practice.

(ii) Weakened relationships between prewhitened residuals: Intuitively, it seems clear than in general the relationships between the residual series $\{c_{jt}\}$ should in some sense be weaker than the relationship among the original $\{z_{jt}\}$, $j = 1,\ldots,k$. This is essentially because when z_{jt} is individually related to its own past values $z_{j(t-1)}, z_{j(t-2)},\ldots$, as in (7.1), these past values serve to a certain extent as proxy variables of past values of other series. (This is indeed one reason why z_{jt} can be forecasted from its own past). Thus, the dynamic relationships between series are partially allowed for by the individual past values so that the structure between the residual series must be weaker. This suggests that it will be more efficacious to model z_t directly than through c_t.

7.2 Econometric models

We have seen earlier in (3.19) and (3.20) that the vector model (3.1) can be regarded as a reduced form of the structural linear simultaneous equation model commonly used in econometric literature. However,

the philosophy which is applied in building "econometric" and "time series" models had been different. On the one hand the structure of econometric models has usually been chosen to reflect economic theory which is believed to apply. On the other hand time series models have usually contained only that structure which was necessary to describe the data. More specifically, the sequential model building process for time series has been directed towards finding a transformation $\pi(B)z_t = \theta^{-1}(B)\phi(B)z_t = a_t$ of the data to white noise uncorrelated with any other known input. Thought of as distinct entities each of these two approaches have points of strength and weakness.

Models derived from theory

1) can be directly related to fundamental mechanism and so, when they provide adequate approximations, can encourage scientific progress; however,

2) economic theory is imperfect and hence certain aspects of such models may be badly wrong; and

3) when there is little information from the data on such questionable aspects, imperfection can go uncorrected.

There has, thus, been a tendency to produce overly elaborate econometric models containing questionable aspects some of which can neither be verified nor discredited by available data.

Empirical models

1) contain only those relationships supported by data; however,

2) they can establish only the existence of complex dynamic correlations which do not necessarily imply direct causation;

3) although they employ a sensible dynamic structure this does not necessarily relate directly to the mechanism; and

4) they necessarily omit external theoretical information which even though not verifiable by available data could be valid and essential to understanding.

In view of this, to the extent that the correlation structure of the system remains constant, empirical time series models can produce good forecasts of future behavior but they may not of themselves explain the mechanism of what is occurring.

The clear conclusion seems to be that the econometric and time series approaches are complementary. The time series approach can indicate inherent dynamic correlative structures in the data which directly or indirectly must be capable of explanation by valid economic theory.

Thus, as in other areas where theory and practice, or model and data, interact, progress not possible by the use of either entity alone may be possible by an iterative conversation between them.

References

[1] Anderson, T.W. (1971). The Statistical Analysis of Time Series, Wiley: New York.

[2] Bartlett, M.S. (1938), "Further aspects of the theory of multiple regression", Proc. Cambr. Phil. Soc., 34, 33.

[3] Box, G.E.B., Erjavec, J., Hunter, W.G. and MacGregor, J.F. (1973), "Some problems associated with the analysis of multiresponse data", Technometrics, 15, 33.

[4] Box, G.E.P. and Cox, D.R. (1964), "An analysis of transformation", J. Roy. Statist. Soc., B, 26, 211.

[5] Box, G.E.P. and Haugh, L. (1977), "Identification of dynamic regression models connecting two time series", J. Amer. Statist. Assoc. 72, 121.

[6] Box, G.E.P., Hillmer, S. and Tiao, G.C. (1978), "Analysis and modelling of seasonal time series", in Seasonal Analysis of Economic Time Series, Department of Commerce, Bureau of the Census ER-1, A. Zellner, ed. 309-33.

[7] Box, G.E.P. and Jenkins, G.M. (1970), Time Series Analysis Forecasting and Control. Holden-Day: San Francisco.

[8] Box, G.E.P. and Newbold, P. (1970), "Some comments on a paper by Coen, Gomme and Kendall", J. Roy. Statist. Soc., A, 134, 229.

[9] Box, G.E.P. and Tiao, G.C. (1975), "Intervention analysis with applications to environmental and economic problems", J. Amer. Statist. Assoc. 70, 70.

[10] Box, G.E.P. and Tiao, G.C. (1977), "A canonical analysis of multiple time series", Biometrika, 64, 355.

[11] Coen, P.G., Gomme, E.D. and Kendall, M.G. (1969), "Lagged relationships in economic forecasting", J. Roy. Statist. Soc. A, 132, 133.

[12] Gray, H.L., Kelley, G.D., and McIntire, D.D., (1978), "A new approach to ARMA modelling", Comm. in Stat. B7, 1.

[13] Hannan, E.J. (1970). Multiple Time Series. Wiley: New York.

[14] Hillmer, S.C. and Tiao, G.C. (1979), "Likelihood function of stationary multiple autoregressive moving average models". To appear in J. Amer. Statist. Assoc.

[15] Quenouille, M.H. (1957). The Analysis of Multiple Time Series. Griffin: London.

[16] Slutsky, E. (1937), "The summation of random causes as the source of cyclic processes", Econometrika, 5, 105.

[17] Tiao, G.C. and Wei, W.S. (1976), "Effect of temporal aggregation on the dynamic relationship of two time series variables", Biometrika 63, 513.

[18] Yule, G.D. (1927), "On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers", Phil. Trans., A226, 267.

Appendix: Data

** MA/SERIES1 **

11.892831	20.476211	14.177125	17.758306	19.929189	10.961964	19.215164	16.384006
16.533008	18.014002	20.699139	13.634472	21.689646	15.427833	18.217524	14.637071
18.471996	13.683318	15.207967	22.437959	18.423503	14.330977	20.713228	16.092812
12.752509	18.990285	15.553543	13.827347	18.562026	14.331301	9.526718	21.504443
18.677275	17.333548	16.475653	15.460480	17.695092	22.306461	15.178916	21.998832
12.255478	15.019931	18.310565	18.781533	17.485580	15.270125	14.429924	17.723277
17.699132	18.402996	13.036323	20.467647	16.578315	18.802513	19.105452	15.626230
17.543900	20.181001	20.654659	16.599173	15.277420	15.710610	19.606784	12.797577
18.797153	15.673999	19.362127	17.299727	17.802382	19.196404	17.180264	17.530818
12.140811	17.626941	16.459136	17.432506	17.881358	15.713799	17.702698	18.623713
19.178451	14.275028	17.196481	17.707756	15.116962	19.155515	20.266668	14.996406
13.616594	15.276367	17.549953	14.335212	19.436909	15.460419	14.796432	17.635208
17.749225	12.548739	19.571478	17.598867	14.385500	16.226130	20.333205	13.956459
20.187557	16.446279	16.250279	14.396787	16.313435	16.647469	20.226595	13.514000
16.576778	17.480394	15.304536	19.798073	16.403132	14.257954	18.765335	18.817010
17.689083	15.251003	13.519662	17.375463	21.831059	15.947061	17.453845	18.703640
15.107664	17.224836	17.531460	18.730959	15.691207	15.096223	20.536828	14.970959
17.547239	18.956430	13.795034	16.614248	11.700365	14.120020	18.040030	18.837307
17.405731	17.940837	17.259035	17.482500	22.853539	14.663418	20.305340	15.321601
16.577325	16.391646	15.015016	20.037449	17.519311	18.035711	16.354772	15.042048
15.999589	16.463719	19.422329	15.629383	17.520103	17.747774	13.857362	19.189862
20.455473	15.464585	17.957288	18.211224	15.089944	17.052711	18.077027	16.754314
14.941989	17.504415	16.231918	17.455851	13.950045	18.025720	16.748367	17.934514
16.052127	22.335728	14.545342	16.721529	18.891933	14.329782	19.709068	14.163682
18.041827	16.387565	16.429702	17.597436	18.448014	15.953017	19.161314	11.766302
18.088107	15.690522	20.340832	19.521496	16.096701	19.599089	16.964158	18.999282
13.940941	19.410328	13.445764	14.727982	14.728442	18.323980	19.843746	21.726571
19.963279	17.562738	19.926497	17.078642	18.312839	17.122561	18.617421	15.379522
18.185206	15.853085	15.555360	18.316904	18.258959	17.555290	21.241123	15.601672
19.535970	16.957357	16.975175	15.375257	15.368377	19.887494	14.756451	18.209120
17.831275	18.549447	18.683672	14.584975	16.034466	19.558124	13.200380	18.233313
14.645736	18.599473						

** MA/SERIES2 **

26.048626	25.546622	22.811276	24.073537	25.814599	24.445776	25.705939	24.082878
22.851145	27.289705	25.879003	25.215001	24.012487	28.242998	22.980079	27.468852
24.243084	25.227159	20.723140	25.718481	27.960523	24.627311	27.172833	25.566029
23.023803	23.175556	26.018130	23.573882	23.403581	24.864253	22.779976	22.983186
23.926095	27.518516	24.458173	26.551991	23.142283	26.640559	26.294783	26.307838
25.900547	23.075784	23.739552	26.510305	26.917736	22.897708	24.029080	24.228299
23.784516	29.691536	21.849575	25.242535	24.969743	26.493911	25.675820	26.341197
24.418432	26.073280	27.840727	26.548908	24.827584	23.020916	26.921408	22.912228
25.676389	23.452591	26.821966	26.492639	24.095465	25.539463	25.758039	27.542534
22.286712	22.605085	24.708892	26.675655	22.867214	26.062911	24.188633	25.756072
26.821031	26.503838	22.867270	24.279416	25.714767	23.378731	28.360101	24.712240
25.976049	22.168530	25.545125	24.103072	24.682615	24.490649	25.414569	23.261413
25.217421	24.804639	23.786081	23.555656	27.043187	23.463145	25.532248	23.610356
26.762499	25.103208	25.592410	23.552294	23.513013	23.755027	28.096668	22.984957
25.170600	24.607644	22.847771	25.783711	27.087705	24.161692	21.740211	27.738531
24.775655	28.250906	20.953649	23.060250	27.104986	25.854233	25.331583	24.956383
28.120575	23.180370	24.578476	25.704901	23.970696	26.058880	24.996382	24.247801
25.579433	26.030440	23.868460	23.587671	24.841186	21.613353	23.410136	25.365785
26.190146	25.680707	25.587964	24.622770	27.688982	24.227477	23.267046	27.697107
25.076285	24.076630	22.524640	26.365212	25.962124	25.673177	25.109259	23.642311
26.225872	23.035877	25.600865	23.271906	26.319722	24.213764	26.561149	23.409723
26.500144	24.799540	24.862906	25.724049	27.416922	21.991016	26.233471	25.619031
23.170909	24.828353	25.753166	26.122762	21.800917	25.896808	23.805124	26.725069
23.474852	25.009526	27.404421	25.253478	23.826616	25.889988	24.706681	25.694834
24.074033	22.522952	27.358077	23.806572	25.286273	26.286319	25.122745	23.928327
24.230844	24.283211	23.708885	27.415904	28.493664	23.850796	27.488586	24.804505
25.877575	25.239878	23.108848	23.733041	23.062104	24.885116	24.898883	26.316015
29.353772	25.768721	27.298434	25.697079	26.614776	25.310586	25.333913	24.961272
25.823218	24.416003	23.418691	25.743477	23.925731	26.720780	27.052003	26.246538
25.128253	25.003986	26.054728	23.988653	24.235702	26.378090	23.717558	25.210339
25.482167	26.615683	25.334724	23.608548	26.050410	25.201568	23.773031	23.989998
24.303349	23.057976						

** AR/SERIES1 **

.818665	5.230414	3.368532	4.369497	6.982769	-.850557	2.299049	3.661262
4.179512	4.402262	8.768697	3.826964	8.181791	4.112360	4.784082	1.008071
3.919423	1.613238	2.297223	9.155984	8.930532	3.893650	6.764202	5.388179
.902125	4.370588	3.197142	.917604	4.680186	2.782328	-3.378120	5.977296
9.873622	8.303306	6.211211	3.495656	5.387010	10.979720	6.749796	9.909164
2.225499	.063328	3.174234	5.726198	6.057990	4.431826	1.662453	3.554870
5.167430	5.932871	2.846239	7.362865	6.752548	7.473092	8.401909	5.013220
4.932010	7.713218	9.747762	6.688653	5.096396	1.750152	4.832703	1.006483
3.722620	3.577487	6.303414	6.664465	7.575537	8.746941	6.702773	5.150693
.025195	2.108001	2.498431	3.493327	5.591526	3.449789	4.231731	5.570420
6.789353	2.633618	3.735401	4.976636	2.502461	5.395088	7.691623	6.129920
.552224	-.227632	2.017074	1.728242	6.219961	5.054211	2.727794	5.071366
6.587389	1.940311	6.466328	7.773634	3.447679	3.468209	8.160621	4.640823
7.280718	6.182211	4.641903	2.345693	3.023676	4.057822	7.897021	4.883103
4.815473	6.684096	5.802826	8.542224	6.825983	3.893897	7.018754	7.771263
7.181672	3.450729	1.215390	3.377686	9.079969	6.684321	5.576341	6.383717
2.735193	4.120496	5.483447	7.116988	4.745922	1.629566	6.357876	4.154614
4.157238	5.268904	2.127995	2.794503	-1.326690	-.150203	5.476046	9.764582
10.185023	10.396656	9.576961	9.053323	13.530212	8.025145		

** AR/SERIES2 **

5.826490	10.094933	8.780906	8.431233	8.103038	4.017174	8.075621	9.682242
8.633091	9.550712	10.895790	7.971834	7.975084	6.144078	5.129679	5.292487
8.524885	10.013151	10.457019	12.335434	10.532402	6.918128	9.237776	9.063469
7.702763	9.285906	9.367434	9.593363	11.653168	11.119566	10.952010	17.248416
16.184350	13.573182	10.786282	10.724560	11.638403	12.972145	9.220657	9.096396
5.189182	5.584249	7.739937	9.426704	10.686816	9.300552	8.287785	9.743738
9.115194	11.680314	10.469078	12.655946	11.254296	11.051714	10.096943	8.221940
8.106866	8.831687	8.911471	6.752504	5.496749	5.060157	8.007108	6.281780
9.210712	8.963438	11.066035	12.235106	11.931648	10.694167	7.808198	7.884865
6.038302	7.245840	7.320343	9.363319	9.103345	7.861016	7.728066	7.673728
7.759572	7.268661	8.401319	8.247339	7.547201	7.766639	8.550904	5.902812
4.198365	4.863707	8.024857	9.550098	12.326655	11.076564	11.261252	12.699358
12.998959	11.776807	14.515920	12.514411	10.758863	11.589024	13.873810	10.508502
11.787075	10.493997	10.198823	9.832729	10.846082	11.089923	14.612116	12.797669
13.974944	15.601188	13.951351	14.057663	13.320208	12.858507	12.360706	11.702556
9.338863	9.587072	9.286538	10.249050	12.392543	9.677939	8.330220	7.236751
7.550385	9.442541	10.163954	10.373633	7.460520	7.019325	9.916505	7.875811
8.065826	9.315096	8.373829	9.371693	10.540309	14.320924	18.605076	20.257987
19.524306	18.578399	17.216337	15.381568	15.801648	11.373024		

** SCC/FT SHARE **

1.3780	1.4061	1.4844	1.4477	1.5003	1.5749	1.6592	1.8158
2.0224	2.1826	2.2902	2.3953	2.5090	2.3415	2.2376	2.2559
2.2144	2.1239	2.2669	2.4675	2.4210	2.0542	1.9759	2.0811
2.2559	2.5800	2.6692	2.8331	3.0250	3.6914	3.9592	3.8406
3.9250	3.8259	3.9580	4.2710	3.7953	3.6168	3.6829	3.4946
3.3222	3.4677	3.6156	3.7819	3.6156	4.1817	4.0974	4.2111
4.4128	4.2368	4.1646	4.1609	3.9531	4.2257	4.2710	4.3982
3.9127	3.6670	3.8944	4.1952	4.4116	4.8640		

** SCC/COMMODITY **

18.7585	18.2769	17.8148	16.8283	16.5690	16.8575	17.5555	18.4719
18.0215	18.0176	17.8694	17.4268	17.7525	17.5263	17.5789	16.9199
17.2455	17.6823	16.7737	16.5436	16.3915	15.9801	15.6038	15.5590
15.6350	15.6233	15.6799	16.1186	16.1000	16.1069	16.0796	15.9606
15.5395	15.1896	15.1320	15.3835	15.5434	15.2236	15.1183	15.0111
14.8668	15.2334	15.6721	15.9450	16.1459	16.5709	16.8283	16.7581
17.6901	18.0196	16.9979	16.6119	16.6586	17.1295	17.5380	17.5867
17.3313	16.2316	15.8339	15.5395	15.3445	15.8900		

** SCC/CAR **

1.0231	1.1790	1.2215	1.4051	1.5514	1.5832	1.7239	1.8110
1.8953	1.9276	2.2020	2.0835	2.2418	2.2421	2.0433	1.7917
1.4744	1.5107	1.4663	2.0672	2.3657	2.4396	2.5780	2.4528
2.5315	2.4128	2.3822	2.4351	2.7650	3.5033	3.6205	3.6352
3.4226	2.4100	1.9905	2.4000	2.7024	2.5519	3.0278	2.9113
3.1288	3.0250	3.5150	3.6595	4.0095	4.4020	4.4539	4.4722
4.2010	4.6038	4.2505	4.1760	3.8615	4.3491	4.2330	4.1271
3.8624	3.2368	3.5601	3.6098	3.6372	3.6925		

** GASFURNACE/GASRATE **

-.109	.000	.178	.339	.373	.441	.461	.348	.127	-.180
-.588	-1.055	-1.421	-1.520	-1.302	-.814	-.475	-.193	.088	.435
.771	.866	.875	.891	.987	1.263	1.775	1.976	1.934	1.866
1.832	1.767	1.608	1.265	.790	.360	.115	.088	.331	.645
.960	1.409	2.670	2.834	2.812	2.483	1.929	1.485	1.214	1.239
1.608	1.905	2.023	1.815	.535	.122	.009	.164	.671	1.019
1.146	1.155	1.112	1.121	1.223	1.257	1.157	.913	.620	.255
-.280	-1.080	-1.551	-1.799	-1.825	-1.456	-.944	-.570	-.431	-.577
-.960	-1.616	-1.875	-1.891	-1.746	-1.474	-1.201	-.927	-.524	.040
.788	.943	.930	1.006	1.137	1.198	1.057	.595	-.080	-.314
-.289	-.153	-.109	-.167	-.255	-.229	-.007	.254	.330	.102
-.423	-1.139	-2.275	-2.594	-2.716	-2.510	-1.790	-1.346	-1.081	-.910
-.876	-.885	-.800	-.544	-.416	-.271	.000	.403	.841	1.285
1.697	1.746	1.683	1.485	.993	.648	.577	.577	.632	.747
.900	.993	.968	.790	.399	-.161	-.553	-.603	-.424	-.194
-.049	.060	.161	.301	.517	.566	.560	.573	.592	.671
.933	1.337	1.460	1.353	.772	.218	-.237	-.714	-1.099	-1.269
-1.175	-.676	.033	.556	.643	.484	.109	-.310	-.697	-1.047
-1.218	-1.183	-.873	-.336	.063	.084	.000	.001	.209	.556
.782	.858	.918	.862	.416	-.336	-.959	-1.813	-2.378	-2.499
-2.473	-2.330	-2.053	-1.739	-1.261	-.569	-.137	-.024	-.050	-.135
-.276	-.534	-.871	-1.243	-1.439	-1.422	-1.175	-.813	-.634	-.582
-.625	-.713	-.848	-1.039	-1.346	-1.628	-1.619	-1.149	-.488	-.160
-.007	-.092	-.620	-1.086	-1.525	-1.858	-2.029	-2.024	-1.961	-1.952
-1.794	-1.302	-1.030	-.918	-.798	-.867	-1.047	-1.123	-.876	-.395
.185	.662	.709	.605	.501	.603	.943	1.223	1.249	.824
.102	.025	.382	.922	1.032	.866	.527	.093	-.458	-.748
-.947	-1.029	-.928	-.645	-.424	-.276	-.158	-.033	.102	.251
.280	.000	-.493	-.759	-.824	-.740	-.528	-.204	.034	.204
.253	.195	.131	.017	-.182	-.262				

** GASFURNACE/CO2 **

53.800	53.600	53.500	53.500	53.400	53.100	52.700	52.400	52.200	52.000
52.000	52.400	53.000	54.000	54.900	56.000	56.800	56.800	56.400	55.700
55.000	54.300	53.200	52.300	51.600	51.200	50.800	50.500	50.000	49.200
48.400	47.900	47.600	47.500	47.500	47.600	48.100	49.000	50.000	51.100
51.800	51.900	51.700	51.200	50.000	48.300	47.000	45.800	45.600	46.000
46.900	47.800	48.200	48.300	47.900	47.200	47.200	48.100	49.400	50.600
51.500	51.600	51.200	50.500	50.100	49.800	49.600	49.400	49.300	49.200
49.300	49.700	50.300	51.300	52.800	54.400	56.000	56.900	57.500	57.300
56.600	56.000	55.400	55.400	56.400	57.200	58.000	58.400	58.400	58.100
57.700	57.000	56.000	54.700	53.200	52.100	51.600	51.000	50.500	50.400
51.000	51.800	52.400	53.000	53.400	53.600	53.700	53.800	53.800	53.800
53.300	53.000	52.900	53.400	54.600	56.400	58.000	59.400	60.200	60.000
59.400	58.400	57.600	56.900	56.400	56.000	55.700	55.300	55.000	54.400
53.700	52.800	51.600	50.600	49.400	48.800	48.500	48.700	49.200	49.800
50.400	50.700	50.900	50.700	50.500	50.400	50.200	50.400	51.200	52.300
53.200	53.900	54.100	54.000	53.600	53.200	53.000	52.800	52.300	51.900
51.600	51.600	51.400	51.200	50.700	50.000	49.400	49.300	49.700	50.600
51.800	53.000	54.000	55.300	55.900	55.900	54.600	53.500	52.400	52.100
52.300	53.000	53.800	54.600	55.400	55.900	55.900	55.200	54.400	53.700
53.600	53.600	53.200	52.500	52.000	51.400	51.000	50.900	52.400	53.500
55.600	58.000	59.500	60.000	60.400	60.500	60.200	59.700	59.000	57.600
56.400	55.200	54.500	54.100	54.100	54.400	55.500	56.200	57.000	57.300
57.400	57.000	56.400	55.900	55.500	55.300	55.200	55.400	56.000	56.500
57.100	57.300	56.800	55.600	55.000	54.100	54.300	55.300	56.400	57.200
57.800	58.300	58.600	58.800	58.800	58.600	58.000	57.400	57.000	56.400
56.300	56.400	56.400	56.000	55.200	54.000	53.000	52.000	51.600	51.600
51.100	50.400	50.000	50.000	52.000	54.000	55.100	54.500	52.800	51.400
50.800	51.200	52.000	52.800	53.800	54.500	54.900	54.900	54.800	54.400
53.700	53.300	52.800	52.600	52.600	53.000	54.300	56.000	57.000	58.000
58.600	58.500	58.300	57.800	57.300	57.000				

** HOUSING/STARTS **

52.1	47.2	82.2	100.9	98.4	97.4	96.5	88.8	80.9	85.8	72.4	61.2
46.6	50.4	83.2	94.3	84.7	79.8	69.1	69.4	59.4	53.5	50.2	39.0
40.2	40.3	66.6	79.8	87.3	87.6	82.3	83.7	78.2	81.7	69.1	47.0
45.2	55.4	79.3	95.0	86.8	81.4	86.4	82.5	80.1	85.6	64.8	53.8
51.3	47.9	71.9	85.0	91.3	82.7	73.5	69.5	71.5	68.0	55.1	42.8
33.4	41.4	61.9	73.8	74.8	83.0	75.5	77.3	76.0	79.4	67.4	69.0
54.9	58.3	91.6	116.0	115.6	116.9	107.7	111.7	102.1	102.9	92.9	80.4
76.2	76.3	111.4	119.8	135.2	131.9	119.1	131.3	120.5	117.0	97.4	73.2
77.1	73.6	105.1	120.5	131.6	114.8	114.7	106.4	84.5	86.0	70.5	46.8
43.3	57.6	77.3	102.3	96.4	99.6	90.9	79.8	73.4	69.5	57.9	41.0
40.0	40.2	62.6	77.9	92.8							

** HOUSING/SOLD **

38.0	44.0	53.0	49.0	54.0	57.0	51.0	58.0	48.0	44.0	42.0	37.0
42.0	43.0	53.0	49.0	49.0	40.0	40.0	36.0	29.0	31.0	26.0	23.0
29.0	32.0	41.0	44.0	49.0	47.0	46.0	47.0	43.0	45.0	34.0	31.0
35.0	43.0	46.0	46.0	43.0	41.0	44.0	47.0	41.0	40.0	32.0	32.0
34.0	40.0	43.0	42.0	43.0	44.0	39.0	40.0	33.0	32.0	31.0	28.0
34.0	29.0	36.0	42.0	43.0	44.0	44.0	48.0	45.0	44.0	40.0	37.0
45.0	49.0	62.0	62.0	58.0	59.0	64.0	62.0	50.0	52.0	50.0	44.0
51.0	56.0	60.0	65.0	64.0	63.0	63.0	72.0	61.0	65.0	51.0	47.0
54.0	58.0	66.0	63.0	64.0	60.0	53.0	52.0	44.0	40.0	36.0	28.0
36.0	42.0	53.0	53.0	55.0	48.0	47.0	43.0	39.0	33.0	30.0	23.0
29.0	33.0	49.0	54.0	56.0							

** INDUSTRY/NEW ORDERS **

293	306	302	336	288	320	301	351	350	336	339	353
349	340	437	428	445	478	407	383	408	387	367	416
402	431	415	419	406	381	349	376	303	301	291	291
341	376	368	392	352	432	370	372	427	331	302	402
337	471	496	455	392	451	399	368	399	376	340	401
422	459	500	484	485	541	419	426	439	438	435	490
491	512	542	597	529	569	526	509	517	526	508	524
533	615	663	657	612	616	559	556	594	613	554	575
628	698	752	702	711	708	625	577	589	634	591	585
574	685	663	732	669	689	610	653	684	589	632	694
694	651	684	760	779	738	649	650	756	731	699	682
726	973	843	878	869	851	757	750	812	797	746	732
846	924	902	878	801	678	706	704	832	736	637	719
598	860	862	860	772	918	724	811	811	809	737	830
864	919	1049	1006	970	1019	905	932	962	963	986	1012
1165	1310	1423	1260	1260	1346	1277	1235	1257	1349	1280	1171
1376	1429	1688	1895	1930	1959	2001	1796	1766	1758	1520	1463
1590	1319	1650	1450	1466	1564	1418	1231	1328	1327	1393	1120
1348	1329	1446	1615	1548	1764	1520	1566	1519	1532	1608	1553
1673	1854	2099	1775	2006	2100	1809	2007	2000	2135	2168	2055
2303	2418	2360	2662	2510	2403	2428					

** INDUSTRY/SHIPMENTS **

312	348	344	357	322	349	310	327	337	302	348	365
320	366	400	425	431	462	399	380	390	380	368	398
363	408	407	441	412	401	372	372	365	335	316	309
312	364	364	379	373	389	305	377	409	356	312	335
305	363	398	445	452	449	383	406	393	382	374	356
337	408	431	474	478	488	414	441	459	453	445	433
428	455	518	549	546	564	506	507	530	503	499	489
482	546	590	604	612	601	561	561	569	571	560	543
559	630	669	685	665	698	612	605	622	628	599	589
551	645	650	719	697	711	623	635	684	595	649	706
620	647	698	734	750	758	694	654	727	714	693	649
660	768	820	831	789	848	746	726	799	769	744	702
787	866	829	820	791	848	730	759	814	723	687	647
759	846	836	824	810	862	718	768	842	772	724	737
846	922	930	978	910	980	890	922	956	935	916	903
1003	1137	1164	1155	1126	1229	1075	1079	1114	1059	1010	993
1096	1225	1324	1338	1311	1493	1334	1380	1554	1621	1556	1463
1492	1689	1670	1665	1592	1693	1490	1526	1535	1595	1518	1419
1449	1647	1664	1682	1693	1784	1614	1635	1755	1641	1595	1581
1660	1880	1944	1948	1965	2060	1838	1687	2019	2033	2032	1988
2087	2220	2278	2435	2337	2490	2305					

** INDUSTRY/TOTAL INVENTORIES **

1233	1232	1228	1193	1166	1146	1124	1107	1096	1091	1083	1087
1099	1136	1152	1156	1176	1190	1197	1188	1187	1191	1188	1222
1249	1303	1318	1299	1279	1275	1229	1198	1179	1167	1165	1147
1180	1143	1138	1132	1119	1111	1093	1074	1052	1048	1059	1086
1116	1146	1174	1169	1162	1151	1147	1146	1152	1166	1178	1212
1242	1249	1256	1236	1217	1203	1190	1181	1185	1184	1198	1246
1260	1266	1286	1286	1283	1289	1285	1301	1316	1345	1387	1436
1482	1480	1490	1493	1487	1493	1506	1517	1529	1542	1564	1604
1643	1669	1681	1689	1708	1737	1744	1752	1820	1858	1911	1940
1959	1992	2008	1993	1975	1953	1923	1909	1909	1911	1957	1978
1976	2005	2024	2048	2042	2030	2000	1993	2019	2007	2037	2089
2108	2169	2180	2196	2212	2204	2212	2231	2253	2281	2336	2399
2413	2444	2483	2510	2537	2548	2560	2551	2514	2508	2563	2630
2645	2688	2716	2730	2737	2771	2767	2759	2738	2737	2891	2925
2806	2788	2762	2711	2722	2694	2664	2661	2679	2713	2809	2990
3027	3055	3092	3124	3159	3165	3186	3226	3264	3321	3418	3541
3676	3819	3933	3986	4072	4165	4285	4386	4489	4599	4764	5095
5234	5433	5594	5690	5621	5659	5613	5595	5562	5560	5593	5598
5620	5755	5819	5785	5718	5674	5628	5592	5555	5561	5681	5764
5709	5812	5710	5636	5519	5467	5413	5428	5473	5492	5565	5893
5959	6056	6119	6162	6190	6315	6376					

** INDUSTRY/UNFILLED ORDERS **

962	920	878	857	823	802	793	817	820	810	809	797
817	837	874	877	891	907	925	928	946	953	952	970
1009	1032	1040	1019	1012	992	969	923	861	827	802	784
813	825	829	882	821	864	889	888	903	878	868	935
967	1175	1131	1139	1079	1061	1077	1039	1045	1039	1005	1059
1135	1186	1255	1205	1272	1325	1330	1315	1295	1280	1270	1327
1390	1497	1471	1519	1501	1516	1526	1519	1506	1529	1538	1573
1624	1693	1766	1819	1819	1834	1862	1857	1882	1924	1918	1950
2028	2096	2180	2197	2203	2253	2266	2234	2205	2211	2203	2199
2222	2222	2235	2249	2220	2197	2148	2202	2202	2196	2179	2167
2241	2245	2231	2267	2296	2250	2205	2201	2230	2247	2253	2296
2352	2557	2590	2627	2707	2740	2751	2775	2788	2818	2820	2850
2900	2967	[illegible]	3094	3104	3134	3114	3050	3077	3100	3040	3112
2947	2961	2967	3003	2965	3021	3027	3070	3039	3074	3189	3182
3203	3197	3316	3344	3404	3443	3458	3468	3479	3499	3569	3678
3880	4013	4272	4377	4511	4628	4830	4896	5129	5518	5888	5966
6186	6350	6710	7267	7086	8352	9019	9055	9087	9784	9748	9749
9986	9376	9450	9281	9115	8946	8921	8620	8422	8154	8029	7731
7833	7315	7097	7313	6805	6865	6771	6699	6461	6359	6367	6339
6352	6326	6461	6308	6347	6389	6500	6499	6461	6503	6697	6766
6082	7173	7312	7551	7732	7609	7169					